


```

QY 67 VYVNDONLDLPEVNELEFYDVSKPGISSYAANPEEAESLIPILKEAENVVPSQOPMT 126
DB 61 VYVNDONLDLPEVNELEFYDVSKPGISSYAANPEEAESLIPILKEAENVVPSQOPMT 120
QY 127 PVYLGATAGRLLEGNAENILOAVDMJ.SNRSALNVOSDAYSILDGTGEGSYLWNTJNY 186
DB 121 PVYLGATAGRLLEGNAENILOAVDMJ.SNRSALNVOSDAYSILDGTGEGSYLWNTJNY 180
QY 187 LIGKLRKRTKTVGVVDLGGGSVQMTYAVSRMTAKNAPVPEGEDEPYIKKLVLOGKKYDL 246
DB 181 LIGKLRKRTKTVGVVDLGGGSVQMTYAVSRMTAKNAPVPEGEDEPYIKKLVLOGKKYDL 240
QY 247 YVHSYLRGRERAEFRFVAGGSANPCILAGFDGAYTYSAGAEYKVSAPASGSNNOCRK 306
DB 241 YVHSYLRGRERAEFRFVAGGSANPCILAGFDGAYTYSAGAEYKVSAPASGSNNOCRK 300
QY 307 IALKALKVNAAPCYONCTFGGIMNGSGSGOKNLFITSSFYLLSDVGIFFVKNPKAKTRP 366
DB 301 IALKALKVNAAPCYONCTFGGIMNGSGSGOKNLFITSSFYLLSDVGIFFVKNPKAKTRP 360
QY 367 VDLKTAACKTNLEDAKSKYPDLKDYKSVYVCLDLYVYVTLVDFGIDPQEVTVVA 426
DB 361 VDLKTAACKTNLEDAKSKYPDLKDYKSVYVCLDLYVYVTLVDFGIDPQEVTVVA 420
QY 427 NEIEYODALVEAAMPGLTAIEAISLSPKPERLMTFYI 462
DB 421 NEIEYODALVEAAMPGLTAIEAISLSPKPERLMTFYI 456

```

RESULT 2

```

Q9FVC2 PRELIMINARY; PRT; 468 AA.
ID 09FVC2:
AC 09FVC2:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE APYRASE GS52.
OS Glycine soja.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3848;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=20496172; PubMed=11043467;
RA Day R.B., McAlvin C.B., Loh J.T., Denny R.L., Wood T.C., Young N.D.,
RA Stacey G.;
RT "Differential expression of two soybean apyrases, one of which is an
RT early nodulin.";
RL Mol. Plant Microbe Interact. 13:1053-1070(2000).
DR EMBL: AF207688; MAG32860.1;
DR InterPro: IPR000407; GDAI_CD39_NTPase.
DR Pfam: PF01150; GDAI_CD39_1.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
SQ SEQUENCE 468 AA; 51817 MW; F9804817D014DB40 CRC64;

```

Query Match 73.0%; Score 1740; DB 10; Length 468;
 Best Local Similarity 73.4%; Pred. No. 2e-124;
 Matches 339; Conservative 51; Mismatches 66; Indels 6; Gaps 5;

```

QY 4 TKAMDPLISLMTFVFM--PAISSOYLGNNTLMNRKILPKNOEPVSYAVTFPAGST 61
DB 10 TKAMDPL-TLFTLLLFTHLSTQYHDGNLTLTHRKIFPK-QEATISYAVTFPAGST 67
QY 62 GSRVHYVNFNDONLDLPEVNELEFYDVSKPGISSYAANPEEAESLIPILKEAENVVPS 121
DB 68 GSRVHYVNFNDONLDLPEVNELEFYDVSKPGISSYAANPEEAESLIPILKEAENVVPS 127
QY 122 QOQNTVPKIGATAGRLLEGNAENILOAVRDMJ.SNRSALNVOSDAYSILDGTGEGSYLM 181
DB 128 LYFTTPVKGATAGRLLEGDAERILQAVRDILKNRSTLMAQPDVAVALITDGTGEGSYLM 187

```

```

QY 182 VTINYLKLGKRRKTVGVVDLGGGSVQMTYAVSRMTAKNAPVPEGEDEPYIKKLVLOG 241
DB 188 VTINYLKLGKRRKTVGVVDLGGGSVQMTYAVSRMTAKNAPVPEGEDEPYIKKLVLOG 247
QY 242 KKYDLYVHSYLRGRERAEFRFVAGGSANPCILAGFDGAYTYSAGAEYKVSAPASGSNL 301
DB 248 KKYDLYVHSYLRGRERAEFRFVAGGSANPCILAGFDGAYTYSAGAEYKVSAPASGSNL 307
QY 302 NOCRRIALKALKVNAAPCYONCTFGGIMNGSGSGOKNLFITSSFYLLSDVGIFFVKNPK 360
DB 308 DEGREVYVQALKNESCPHONCTFGGIMNGSGSGOKNLFITSSFYLLSDVGIFFVKNPK 367
QY 361 NAKRPVYDLKTAACKTNLEDAKSKYPDLKDYKSVYVCLDLYVYVTLVDFGIDPQ 420
DB 368 NSKIHVPYDLKTAACKTNLEDAKSKYPDLKDYKSVYVCLDLYVYVTLVDFGIDPQ 426
QY 421 QEVTVANEIEYODALVEAAMPGLTAIEAISLSPKPERLMTFYI 462
DB 427 QEVTVANEIEYODALVEAAMPGLTAIEAISLSPKPERLMTFYI 468

```

RESULT 3

```

ID 09A015 PRELIMINARY; PRT; 466 AA.
AC 09A015:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE APYRASE.
GN APY1.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
OX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=21196099; PubMed=11299390;
RA Cohn J.R., Uhm T., Ramu S., Nam Y.W., Kim D.J., Penmetla R.V.,
RA Wood T.C., Denny R.L., Young N.D., Cook D.R., Stacey G.;
RT "Differential regulation of a family of apyrase genes from Medicago
RT truncatula.";
RL Plant Physiol. 125:2104-2119(2001).
DR EMBL: AF288132; AAK15160.1.
DR InterPro: IPR000407; GDAI_CD39_NTPase.
DR Pfam: PF01150; GDAI_CD39_1.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
SQ SEQUENCE 466 AA; 51561 MW; 122A9004A2D882DF CRC64;

```

Query Match 71.8%; Score 1710; DB 10; Length 466;
 Best Local Similarity 69.7%; Pred. No. 3.8e-122;
 Matches 324; Conservative 60; Mismatches 77; Indels 4; Gaps 3;

```

QY 1 LHM--TKAMDPLISLMTFVFM--PAISSOYLGNNTLMNRKILPKNOEPVSYAVTFPAGST 58
DB 3 LHMQTKNMNMNMTLITLFLFIMPISYSQYLGNNTLMNRKIPK-OEPISSIAVYVDA 61
QY 59 GSTGSRVHYVNFNDONLDLPEVNELEFYDVSKPGISSYAANPEEAESLIPILKEAENVV 118
DB 62 GSTGSRVHYVNFNDONLDLPEVNELEFYDVSKPGISSYAANPEEAESLIPILKEAENVV 121
QY 119 PVSQOPNTPVKIGATAGRLLEGNAENILOAVRDMJ.SNRSALNVOSDAYSILDGTGEGS 178
DB 122 PEDQSKTPIRKIGATAGRLNGDASEKILQSVHDSNRSSTFNVDPAVSIIDGTGEGC 181
QY 179 YLWTVINYLKLGKRRKTVGVVDLGGGSVQMTYAVSRMTAKNAPVPEGEDEPYIKKLV 238
DB 182 YLWTVINYLKLGKRRKTVGVVDLGGGSVQMTYAVSRMTAKNAPVPEGEDEPYIKKLV 241
QY 239 LOGKKYDLYVHSYLRGRERAEFRFVAGGSANPCILAGFDGAYTYSAGAEYKVSAPASG 298

```

Query Match	70.7%	Score 1684	DB 10	Length 462
Best Local Similarity	68.5%	Pred. No. 3,6e-120		
Matches 319	Conservative 68	Mismatches 71	Indels 8	Gaps 5
QY	1	LHW---	TKADELISLMTFVFMILPAISSQYIGNNILMNKILLPKNOEPTSYAVIF	56
Db	1	MMWVPKPKRSFLL-LTFELFLSPKSSSQYVNSILNHRKILP-NQELLTSYAVIF	58	
QY	57	DAGTSGSVHYHNPQONDLIPVEMLEFYDSVKRGLSSYANPEAASLLPILKEAN	116	
Db	59	DAGSSGRVHVFNFQONDLHIGNDLEFKTKIKPGLSSYAKPEAAASLLPLLEEAD	118	
QY	117	VVPVSOQNPPTPKVLGATAGLRLLBENAAENIIQAVRDMLSNRSALNVQSDAYSILDTGE	176	
Db	119	VYPELHPTKPTPLKLGATAGLRLLDDGAAPKILQAVREMRNRSLSIQDVAVSIDTGE	178	
QY	177	GSLYMTVNTYLIGRKGRFTKTVGVVDLGGGSVQMTYAVSRNTAKNAPVPEGEDPYIKK	236	
Db	179	GSLYMTVNTYLIGRKGRFTKTVGVYIDLGAGVQMAVAVSRNTAKNAPVPEGEDPYMKK	238	
QY	237	LVLOQKKYDLYHSYLARKRGREAFRAEIFEYVAGSASNPCLIAFGDGYATYSGAEYKVSAPA	296	
Db	239	LYLKAKKTDLYHSYLARKRGNDAAARKIFPTTDGASPCLLAGTEDYLRKSGESYNIYGT	298	
QY	297	SGSNINQCRKIALKALKVNAACPYNQCTFFGILWNGGSGGOKNLFTLSFYLSDEVGIF	356	

Db	299	SGANNECRDLAQLRLNPEPCSHENCFCGGIMDGKSGQKLNVLVTSAFYRSSEVG-F	357
Qy	357	VNKNPAKTRPVDLTKAAKLACTNIEDAKSKYDLYEKDSYVCLDLVYTTLLVDGFG	416
Db	358	VTTPSKNRKPLDFTFAACQACSLTFEEFKSTFPPN-EKDKLPFCVDEFTYOTLLVDGFG	416
Qy	417	LDPPEVTVANIEYODALVEAMPLGRTAIEAISLSPFERLMYFI	462
Db	417	LDPPEVTVANIEYODALVEAMPLGRTAIEAISLSPFERLMYFI	462
RESULT	5		
Q9FEA6	Q9FEA6	PRELIMINARY:	PRT: 455 AA.
AC	Q9FEA6:		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	APYRASE [EC 3.6.1.5] (NUCLEOSIDE TRIPHOSPHATASE) (APYRASE S-TYPE) (S-TYPE APYRAS PRECURSOR).		
OS	NPYASE.		
OS	Pisum sativum (Garden pea).		
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
CC	eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.		
OX	NCBI_TaxID=3888;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ALASKA; TISSUE=DARK GROWN STEM INTERNODE;		
RA	Shibata K., Abe S., Davies E.;		
RT	"Structure of the coding region and mRNA variants of the apyrase gene		
RT	from pea (Pisum sativum).";		
RL	Acta Physiol. Plant. 0:0-0(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ALASKA; TISSUE=DARK GROWN STEM INTERNODE;		
RA	Shibata K., Abe S., Davies E.;		
RT	"Structure of the coding region and mRNA variants of the apyrase gene		
RT	from pea (Pisum sativum).";		
RL	Acta Physiol. Plant. 0:0-0(2001).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ALASKA; TISSUE=DARK GROWN STEM INTERNODE;		
RA	Shibata K., Morita Y., Abe S., Stankovic B., Davies E.;		
RT	"Apyrase from pea stems: isolation, purification, characterization and		
RT	identification of a NTPase from the cytoskeleton fraction of pea stem		
RT	tissue.";		
RL	Plant Physiol. Biochem. 37:881-888(1999).		
DR	EMBL; AB038666; BAB18895.1; -		
DR	EMBL; AB032621; BAB18890.1; -		
DR	EMBL; AB038554; BAB18893.1; -		
DR	EMBL; AB027613; BAB40230.1; -		
DR	Interpro: IPR000407; GDAI_CD39_1_NTPase.		
DR	Pfam; PF01150; GDAI_CD39_1.		
DR	PROSITE; PS01238; GDAI_CD39_NTPASE; 1.		
KW	Hydrolase; Translit peptide.		
FT	TRANSIT		
FT	1		
FT	20		
FT	POTENTIAL.		
SEQUENCE	455 AA; 50106 MM; 57335F2FC576111 CRC64;		
Query Match	70.6%; Score 1681.5; DB 10; Length 455;		
Best Local Similarity	68.9%; Pred. No. 5.5e-120;		
Matches 315; Conservative 64; Mismatches 75; Indels 3; Gaps			
Qy	7	MDFLSLMTFVFMMPAIISSSOYLGNNTILMKKILLPKNOEYVTSYAVTFDAGSTGRVH	66
Db	1	MEFLKLTFTFLFLESPATISSQYLGNNTLTSKRIFL--KOEIISSYAVVFDAGSTGRSH	58
Qy	67	VYNFOONDLPLVEVNELEDFYDVKRGSLSSYANPEEAESLLPLLKEAENVVPSQOPT	126
Db	59	VYHFNQNDLILHTIGGVETYNKTTGTGLSSYANPEQAASLLPLLEQAEVDVYPDLOKRT	118
Qy	127	PVKIGATAGLRLLGEGNAENIIQAAVRDMLSNRSALNVQSDAVSILDGTQEGSYLWVTNY	186

```

119 PVRIGATAGLRRLINGDASEKIIQSVADMLSNRSTFNVPDANISIDGQEGSYLWATYNY 178
187 LGKLGKREFTKTVGVYDGGGVOMTYAVSRNTANAPVEGEPYIKKLVLQKKYDL 246
179 ALGNLCKRTTKTVGVYDGGGVOMAYAVSKRTANAPVADGDDPYIKKVLKCTPYDL 238
247 VVHSLRYGREAFAEIRFVAGSANPCIIAGFDGATYTSAGEYKVSAPASGNLNOCRK 306
239 VVHSLYHFGREASRAEIIKLTPRSPNCPCLAGFNQIYISGEEFATATYTSAGANNNCKN 298
307 IALKALKVAPCPYQNCFTGGIWMGGSGGKNLFTSSFYLLSDVG-IFVKNPNAKIR 365
299 TTRKALKVAPCPYQNCFTGGIWMGGSGGKNLFTSSFYLLSDVG-IFVKNPNAKIR 358
366 PVDLTKAAKIACKTNLEDAKSKYPDLYEKDSVEYVCLDLYVYVTLVNGFGLDPQEVTV 425
359 PVDITKAKKEACALNEFDKSTYPLDKKNVASYVCMQDIYQYVLVDGFLDPLQKITS 418
426 ANEIEYODALVEAAMPGLGTAIEAIISSLPKPERLMYFI 462
419 GKEIEYODALVEAAMPGLGTAIEAIISSLPKPERLMYFI 455

```

```

RESULT 6
ID O9SPM6 PRELIMINARY; PRT; 455 AA.
AC O9SPM6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE NOD FACTOR BINDING LECTIN-NUCLEOTIDE PHOSPHOHYDROLASE.
GN LNP.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA MEDLINE=99444909; PubMed=10517321;
RA Roberts N.J., Brigham J., Wu B., Murphy J.B., Volpin H.,
RA Phillips D.A., Etzler M.E.;
RT "A Nod factor-binding lectin is a member of a distinct class of
RT apyrases that may be unique to the legumes.";
RL Mol. Gen. Genet. 262:261-267(1999).
DR EMBL: AF156782; AAF00611.1;
DR InterPro: IPR000407; GDAI_CD39_NTPase.
DR Pfam: PF01150; GDAI_CD39; 1.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
KW Lactin; Hydrolase.
SQ SEQUENCE 455 AA; 49879 MW; 68122846D7EC361B CRC64;

```

```

Query Match 69.8%; Score 1662.5; DB 10; Length 455;
Best local Similarity 68.8%; Pred. No. 1.5e-118;
Matches 315; Conservative 66; Mismatches 72; Indels 5; Gaps 4;
QY 7 MDTLISLM-IFVMLMPAIISSOYLGNMILNKRILLPKNOEPTVSYAVIFDAGSTGSHV 65
DB 1 MERLTLIAIFLLMLPAITSSOYLGNMILNKRIF--QKQETLTSYAVIFDAGSTGSHV 58
QY 66 HYNFPOONDLLEVENLEFYDSVKPGLSSYANPEEAESLIPLKEANVPVVSQON 125
DB 59 HYNFPOONDLLEHNDIEFDVKIPGLSAYGNPEQAKSLIPLEEAEDVVPEDLHK 118
QY 126 TVYKLGATAGLRRLINGDAENIIQAVADMLSNSALNVOGDAVSIIDGTOEGSYLWATN 185
DB 119 TPLRLGATAGLRRLINGDAEKILQATRNMFNSRSTLNVOGDAVSIIDGTOEGSYLWATN 178
QY 186 YILGKLGKREFTKTVGVYDGGGVOMTYAVSRNTANAPVEGEPYIKKLVLQKKYD 245

```

```

DB 179 YVLNIGKSFTEKSVGVIDLGGGVOMTYAVSKRTANAPVADGDDPYIKKLVLQKKYD 238
QY 246 LVHSLRYGREAFAEIRFVAGSANPCIIAGFDGATYTSAGEYKVSAPASGNLNOCR 305
DB 239 LVHSLYHFGREARAQVLANATNSANPCIIIPGNGFTTSGYEYKAFSSSSNDDCK 298
QY 306 KIALKALKVAPCPYQNCFTGGIWMGGSGGKNLFTSSFYLLSDVG-IFVKNPNAKI 364
DB 299 EILKVLKLVNDPCYPSCFTGGIWMGGSGGKNLFTSSFYLLSDVG-IFVKNPNAKI 358
QY 365 RPVDLTKAAKIACKTNLEDAKSKYPDLYEKDSVEYVCLDLYVYVTLVNGFGLDPQEVTV 424
DB 359 HPVDFELKARACALNEFDKSTYPLTDKAR-PYVCMQDIYQYVLVDGFLDPLQKITS 417
QY 425 VANEIEYODALVEAAMPGLGTAIEAIISSLPKPERLMYFI 462
DB 418 VGEIGIYQNSVBEAAMPGLGTAIEAIISSLPKPERLMYFI 455

```

```

RESULT 7
ID O9SLV4 PRELIMINARY; PRT; 447 AA.
AC O9SLV4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE APYRASE (EC 3.6.1.5) (FRAGMENT).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALASKA; TISSUE=DARK GROWN STEM INTERNODE;
RA Shibata K., Abe S.;
RT "Isolation and identification of a cytoskeleton-associated nucleotide
RT triphosphatase in pea stem cells.";
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB027616; BAA89275.1;
DR InterPro: IPR000407; GDAI_CD39_NTPase.
DR Pfam: PF01150; GDAI_CD39; 1.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
KW Hydrolase.
FT NON_TER
SQ SEQUENCE 447 AA; 49065 MW; 7D0F21B2AA94BF5D CRC64;

```

```

Query Match 69.2%; Score 1647.5; DB 10; Length 447;
Best local Similarity 68.8%; Pred. No. 2.1e-117;
Matches 309; Conservative 62; Mismatches 75; Indels 3; Gaps 2;
QY 15 TFVFLMMPAIISSOYLGNMILNKRILLPKNOEPTVSYAVIFDAGSTGSHVYNFQONL 74
DB 1 TPLFLMMPAIISSOYLGNMILNKRILLPKNOEPTVSYAVIFDAGSTGSHVYNFQONL 58
QY 75 DLLPENLEFEYDSVKPGLSSYANPEEAESLIPLKEANVPVVSQONTPKLGATA 134
DB 59 DLLHIGKGVYENKTKTTPGLSSYANPEEAESLIPLEQADVDVDDQKTPRLGATA 118
QY 135 GLRLGDAENIIQAVADMLSNSALNVOGDAVSIIDGTOEGSYLWATNLYLGLKGR 194
DB 119 GLRLNGDASEKIIQSVADMLSNRSTFNVPDANISIDGQEGSYLWATNLYLGLKGR 178
QY 195 FTKTVGVYDGGGVOMTYAVSRNTANAPVEGEPYIKKLVLQKKYDLYVHSTLYR 254
DB 179 YTKTVGVYDGGGVOMAYAVSKRTANAPVADGDDPYIKKVLKCTPYDLVHSTLYR 238
QY 255 GREAPRAEIRFVAGSANPCIIAGFDGATYTSAGEYKVSAPASGNLNOCRKIALKALKY 314
DB 239 GREASRAEIRFVAGSANPCIIAGFDGATYTSAGEYKVSAPASGNLNOCRKIALKALKY 298
QY 315 NACPYPQNCFTGGIWMGGSGGKNLFTSSFYLLSDVG-IFVKNPNAKIRPVDLTKAA 373

```


[illegible]

SEQUENCE	407 AA	44577 MW	644D3618473059545 CRC64
09AVN8	PRELIMINARY	PRT	407 AA.
AC	09AVN8		
AD	01-JUN-2001 (TREMblrel. 17, Created)		
AE	01-JUN-2001 (TREMblrel. 17, Last sequence update)		
AF	01-OCT-2001 (TREMblrel. 18, Last annotation update)		
AG	S-TYPE APPRASE (EC 3.6.1.5) (FRAGMENT).		
AH	Pisum sativum (Garden pea).		
AI	Eukaryota: Viridiplantae; Streptophyta; Tracheophyta;		
AJ	Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots: Rosidae;		
AK	eucots II: Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.		
AL	NCBI_TaxID=3888;		
AM	[1]		
AN	SEQUENCE FROM N.A.		
AO	STRAIN-ALASKA; TISSUE-DARK GROWN STEM INTERNODE;		
AP	Shibata K., Abe S., Davies E.;		
AQ	"Structure of the coding region and mRNA variants of the asprase gene		
AR	from pea (Pisum sativum).";		
AS	Acta Physiol. Plant. 0:0-0(2001).		
AT	[2]		
AV	SEQUENCE FROM N.A.		
AW	STRAIN-ALASKA; TISSUE-DARK GROWN STEM INTERNODE;		
AX	Shibata K., Morita Y., Abe S., Stankovic B., Davies E.;		
AY	"Asprase from pea stems: Isolation, purification, characterization and		
AZ	identification of a NTPase from the cytoskeleton fraction of pea stem		
BA	tissue.";		
BB	Plant Physiol. Biochem. 37:881-888(1999).		
BC	EMBL; AB027615; BAB40231.1;		
BD	InterPro: IPR000407; GDAI_CD39_NTPase.		
BE	Pfam: PF01150; GDAI_CD39; 1.		
BF	PROSITE; PS01238; GDAI_CD39_NTPASE; 1.		
BG	Hydrolase.		
BH	NON_TER		
BI	1		
BJ	1		
BK	SEQUENCE		

	Query Match	64.28	Score 1528.5	DB 10	Length 407
	Best Local Similarity	69.88	Pred. No. 2.2e+08		
	Matches	284	Conservative	55	Mismatches 67; Indels 1; Gaps
Oy	57	DAGTSGRVHVVNFDDILLPVENLEEFYDSVGRGLSSYAANEAEASLIPLKEAEN	116		
		: : : : :			
Dd	1	DAGSTGRIHYHHNOMIDLHIKRGVEYYNRKIPGSLSYANNEDEAKSIIPLLEQED	60		
Oy	117	VVPVSOCNFEVKGATGADGLELGGNAENTLAQVRMLNSRSLANTOSDASTIDTOE	176		
		: : : : :			
Dd	61	VVPDDLPQRKTPVRIGATRLGRLFLNGDASEKLTSVROMLSRSFTFNVPQDAVSIIDTOE	1200		
Oy	177	GSIYLVITINYLGLKLGKRFRTKVGVVDLGGGSVOMTAAVBNTKANAPKYECGEDPYIKK	2368		
		: : : : :			
Dd	121	GSYLVIVTNVAALGNIGKKRYTTAVVIDLGGGSYOMAABVASKTKPNPKAADGDPPYIKK	1800		
Oy	237	LVLGGKKRDLLVHSHLYRFGREAFEAIEFKVAGSGSANCIILAGFCAGTAYGAEEKAVAPA	2666		
		: : : : : : : : : : : : : : : : : : : :			
Dd	181	VVLKGIPIPDYLVHSYHLHREGRAESAELIKLTPRSPNPCLLAGFNGIITYTSEEKKAIAVT	2400		
Oy	297	SGSNLNOCRKIALAKALKYNACPCFPQNCTFGITNNWGGSGSGKNLFUTSRFYLSBDVG-I	3555		
		: : : : :			
Dd	241	SGAANKNKNTIRAKLKNTNYCPQNCTFGGITNNWGGGNGCKNLTEASSFFPYLPEDIMGV	3000		

```

OY      356  FVNNKNNAIRPDDTAKLACKNLDEAKSKYDYLXKDSVEVCCDLVYVTLVDGF  415
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      301  DASTPNEILRPVDIEITKKNKCALNFBDAKSTYFFLDKKNNAASYVCHDILQIYLLVDGF  360
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
OY      416  GIDPQEVTVANEIYQDALVEAMPUGCTALEATSSLPKFFRLMYFI  462
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      361  GIDPQKTTITSGKEIYQDAIYEAMPUGNAVEATSLPKFFRLMYFV  407
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

Q9SPM7	9	PRELIMINARY;	PRT;	467 AA.
ID	Q9SPM7			
AC	Q9SPM7;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 13, Last annotation update)			
DE	APYRASE.			
OS	Dolichos biflorus (Horse gram).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eucotyledons I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.			
OX	NCBI_taxid=3840;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99444909; PubMed=10517321;			
RA	Roberts N.J., Brigham J., Wu B., Murphy J.B., Volpin H.,			
RA	Phillips D.A., Etzler M.E.;			
RT	"A Nod factor-binding lectin is a member of a distinct class of			
RT	apyrases that may be unique to the legumes.";			
RL	Mol. Gen. Genet. 262:261-267(1999).			
DR	EMBL; AF156781; AAF00610.1; -			
DR	InterPro; IPR000407; GDAL_CDS9_NTPase.			
DR	Pfam; PF01150; GDAL_CDS9; 1.			
DR	PROSITE; PS01238; GDAL_CDS9_NTPASE; 1.			
SO	SEQUENCE 467 AA; 51164 MW; 8FCC200AA60D7376 CRC64;			

[illegible]

```

RESULT 10
ID 09FVC3 PRELIMINARY: PRT: 463 AA.
AC 09FVC3:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE APYRASE G550 (FRAGMENT).
OS Glycine soja.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20496172; PubMed=11043467;
RA Day R.B., McAlvin C.B., Loh J.T., Denny R.L., Wood T.C., Young N.D.,
RA Stacey G.;
RT "Differential expression of two soybean apyrases, one of which is an
RT early nodulin."
RL Mol. Plant Microbe Interact. 13:1053-1070(2000).
DR EMBL: AF207687; AGS2959.1; -.
DR InterPro: IPR000407; GDAI_CD39_NTPase.
DR Pfam: PF01150; GDAI_CD39_1.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
DR NON_TER 463 463
SO SEQUENCE 463 AA; 50435 MW; FCB5B8730F7669 CRC64;

Query Match 59.0%; Score 1404.5; DB 10; Length 463;
Best Local Similarity 60.6%; Pred. No. 7,7e-99;
Matches 282; Conservative 62; Mismatches 106; Indels 15; Gaps 7;

QY 8 DELSLMTFVFMIPAISSSQ-----YLGNIIIMNRKILPKNOEPT---SYAVIPDAG 59
DB 4 DYLLILSLSTVLAASSSSSSSSSYLVLGK--FSHRLKLSPIYIIHKTIDESTYAVITDAG 61
QY 60 STGSRVHYVNDQDLIPENELFEYDSVPGCLSTYAANDEAEASLIPLEKEANVVP 119
DB 62 STGSRVHYVNRNODLIRIGODELFVKMPGISAAYENPODAESLIPLEAEAAVP 121
QY 120 VSOQNTPVKIGATAGRLLEGNAENILQAVRDMLSNRSALNVSQAVSILDTGSGSY 179
DB 122 QEFHPRTPVKIGATAGRLLEGDAEDRILOAVSDMLKNRSTLVNGADAVSLGNDGAY 181
QY 180 LMTVINLLGKLRKFTKYGVVDLGGGSVQMTYAVSRNTAKNAKPVYEGEDPYIKLYL 239
DB 182 QMTVINLLGNLGRKHYSETAVVVDLGGGSVQMAVAVSETDAKAPRAPDGVESYITEMFL 241
QY 240 QGKKYDYVHSYLRGEAFRAELFEYKAVAGSANPCILAGDGAITYSGAEYKVASPASGS 299
DB 242 RGGKYYLYVHSYLRGELARAEALKVR-HSENPCLLAGDGYVVGQYKAKAPSPGS 300
QY 300 NLNOCRKIALKALKVNAPCYPQONCTFGGIMNGGSGGOKNLFITSSPFYLSDEVGTFV-- 357
DB 301 SFGSCQNVVVEALHVNATGCTYKDCFTFGIMNGGSGAGENNPFITASFEEYVADAG-FVDP 359
QY 358 NKNNAKRPVDLTAAKLACTNLEDAKSKYPDLYEKDSVEYVCLDLYVYVTLVDGFGI 417
DB 360 NAENAKRPVDFENAKVACNTLKDLSIFPRVKDGD-VPICLDLYVEYVTLVDGFGI 418
QY 418 DPFQEVYVANEIEYODALVEAAMPILGTATATATATATATATATATATATATATATAT 462
DB 419 DPFQEVYVANEIEYODALVEAAMPILGTATATATATATATATATATATATATATATAT 463

```

```

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE APYRASE 2.
GN Apy2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3886;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin L., Roux S.J.;
RT "Cloning of a Second Apyrase in Pisum sativum."
RL Submitted (SEP-2000) for the EMBL/GenBank/DBJ databases.
DR EMBL: AF305783; AAG22044.1; -.
DR InterPro: IPR000407; GDAI_CD39_NTPase.
DR Pfam: PF01150; GDAI_CD39_1.
DR PROSITE: PS01238; GDAI_CD39_NTPASE; 1.
SO SEQUENCE 447 AA; 48783 MW; 5CB2B0CC7A7F660E CRC64;

Query Match 57.3%; Score 1365.5; DB 10; Length 447;
Best Local Similarity 59.0%; Pred. No. 6,9e-96;
Matches 261; Conservative 83; Mismatches 93; Indels 5; Gaps 4;

QY 13 LMTFVFMIPAISSSQYLGNNIIMNRKILPKNOEPTSYAVIPDAGSTGSRVHYVNDQ 72
DB 8 LVTLLILYIMPTSHSIEDYALTHKTI--SPDKLSDSFAYIFDGSGSRVHYVFRDR 65
QY 73 NDDLPEVNELEFYDSVKGELSSYAANPEAEASLIPLEKEANVVPVSOQNTPVKIG 132
DB 66 NDLVHIGNDLEFVLQIKPGLSAQDPQAAESLILDKAESVYPMERYMTPVAVGA 125
QY 133 TAGRLLEGNAENIILQAVRDMLSNRSALNVSQAVSILDTGSGSYLMTVINLLGKIG 192
DB 126 TAGRLLEGDAEDRILOAVRDLKORSTLKSFGAAVAVLDGQEGAFQVYVTLNLLGKIG 185
QY 193 KREFTVGVVDLGGGSVQMTYAVSRNTAKNAKPVYEGEDPYIKLYLQGGKDYVHSYL 252
DB 186 RVDSDYGVVDLGGGSVQMAVAVISSETSAKAPKVPDGPYKEMFNGRKYYLVHSL 245
QY 253 RYGRERFAREIFRYVAGSANPCILAGDGAITYSGAEYKVASPASGSNLNOCRKIALKAL 312
DB 246 RYGLLAAREELIKISDSDSNPCILAGYDSYVYGKSEFKAASSPSGSSLINECKGIALKTL 305
QY 313 KYN-APCPYONCTFGGIMNGGSGGOKNLFITSSPFYLSDEVGTFVNR-PNAKIRPVDLK 370
DB 306 KYNESICTHMKCTFGGIMNGGSGGDKNLFARSFFFYRAEAGFADPKSPAKVAVPDPFE 365
QY 371 TAAKLACTNLEDAKSKIPDLYEKDSVEYVCLDLYVYVTLVDGFGIDPFQEVYVANEIE 430
DB 366 KAAKLACQPKLEDAKSTYPNV-EEGNLPIYCLMDLYQYTLVFGFGIDPMOQITLVKGVK 424
QY 431 YODALVEAAMPILGTATATATATATATATATATATATATATATATATATATATAT 452
DB 425 YHDSLVEAAMPILGTATATATATATATATATATATATATATATATATATATATAT 446

RESULT 12
ID 09SPM5 PRELIMINARY: PRT: 472 AA.
AC 09SPM5:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE APYRASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
SO STRAIN=CV. NOSSEN (NO-0);

```

RX MEDLINE-99444909; PubMed-10517321;
 RA Roberts N.J., Brigham J., Wu B., Murphy J.B., Volpin H.,
 RA Phillips D.A., Etzler M.E.;
 RT "A Mod factor-binding lectin is a member of a distinct class of
 RT apyrases that may be unique to the legumes.";
 RL Mol. Gen. Genet. 262:261-267(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-98403884; PubMed-9734815;
 RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT physically assigned pl and TAC clones.";
 RL DNA Res. 5:203-216(1998).
 DR EMBL, AF156783; AAF00612.1; -;
 DR EMBL, AB012246; BAB09486.1; -;
 DR InterPro: IPR000407; GDA1_CD39_NTPase.
 DR Pfam: PF01150; GDA1_CD39; 1.
 DR PROSITE: PS01238; GDA1_CD39_NTPase; 1.
 SQ SEQUENCE 472 AA; 51599 MW; FA0C280456D51224 CRC64;

Query Match 52.1%; Score 1241; DB 10; Length 472;
 Best Local Similarity 55.2%; Pred. No. 2.4e-86;
 Matches 245; Conservative 81; Mismatches 112; Indels 6; Gaps 5;

QY 11 ISLMFVFMALPAISSQYLGNILMNRKILLPKNOEPTSTAVIFDAGSTGRVHYNF 70
 DB 31 IYLIAGVLLMLPGRSISDSVEEYSVHNRRGGPNSRGP-KNYAVIFDAGSGSRVHYCF 89
 QY 71 DONLDLIPVLENLEFYDSYKPGLSAANPEEAESLIPLKEAENVVVSQOQNPVTL 130
 DB 90 DONLDLIPVLENLEFLQKPGLSAAYPTDPOANSVSLDKAENVPRELRPKTHVY 149
 QY 131 GATAGRLLEGNAENIIQAVRDMLSNRSALNVQSDAVSILDGTQGSYLTMTINLLGK 190
 DB 150 GATAGRTIGHASENIIQAVRELRDRMLKTENAAVYLDGTQGSQWTTINLLKN 209
 QY 191 LCKRTFTYGVVDLGGGSQVMTYAVSRNTAKNAKPVPEGEDPYIKKVLQKKYDLVHS 250
 DB 210 LCKPYSDFGVVDLGGGSQVMAVISEEDAAAPKPLEGEDSVREMYLGRKRYFLVHS 269
 QY 251 YLRVREARERATFEFYVAGGSANPCILAGDGAITYSGAITYKSAPASGSLNOCRIALK 310
 DB 270 YLHYGLARARIELKVSSEDSNPCTAGTYKTKGKAFAKASPSGASLDECRITTN 329
 QY 311 ALKVN-APCPYONCTFGGIMNGGGSGQKNLFTSSFYLSFVLSDEVGIFVN--KPNAKIRPV 367
 DB 330 ALKVNDTLCTHMKCTFGGVWNGGGGQKNMFVASFPPRAAEAG-FVDPNQPVAIVRSM 388
 QY 368 DLKTAAKLACKTNLEDAKSKYIPDLYEKDSVEYVCLDLYVYLLVDGFGLDPPQEVTVAN 427
 DB 389 DEKKAARAKCSKMLEGKSTFP-LVEEENLPYLCDLVYQYTLILDGFGLEPSQITLVK 447
 QY 428 EIEYODALVEAAMPGLGTAIEAISS 451
 DB 448 KYKYGDAVEAAMPGLGSAIEAVSS 471

RESULT 13
 ID 09SGG2 PRELIMINARY; PRT: 471 AA.
 AC 09SGG2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE APYRASE.
 GN ATAPY1 OR T6K12.30.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID-3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jeter C.R., Thomas C.E., Roux S.J.;
 RT "Cloning and sequencing of cDNA homologous to pea apyrase.";
 RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Romling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC T6K12 genomic sequence.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF093604; AAF00071.1; -;
 DR EMBL, AC016829; AAF26805.1; -;
 DR InterPro: IPR000407; GDA1_CD39_NTPase.
 DR Pfam: PF01150; GDA1_CD39; 1.
 DR PROSITE: PS01238; GDA1_CD39_NTPase; 1.
 SQ SEQUENCE 471 AA; 51193 MW; 6EF17A73207ECD42 CRC64;

Query Match 52.1%; Score 1240.5; DB 10; Length 471;
 Best Local Similarity 55.1%; Pred. No. 2.6e-86;
 Matches 247; Conservative 74; Mismatches 112; Indels 15; Gaps 6;

QY 11 ISLMFVFMALPAISS---QYLGNNILMNRKILLPKNOEPTSTAVIFDAGSTGRVHY 66
 DB 31 IYLIAGVLLMLPGRSISDSVEEYSVHNRRGGPNSRGPKN-----YAVIFDAGSGSRVHY 84
 QY 67 YVNFQDNLDLIPVLENLEFYDSYKPGLSAANPEEAESLIPLKEAENVVVSQOQNPVTL 126
 DB 85 YVCFQDNLDLIPVLENLEFLQKPGLSAAYPTDPOANSVSLDKAENVPRELRPKTHVY 144
 QY 127 PVKLGATAGRLLEGNAENIIQAVRDMLSNRSALNVQSDAVSILDGTQGSYLTMTINYL 186
 DB 145 PVKVGATAGRLGHASENIIQAVRELRDRMLKTENAAVYLDGTQGSQWTTINLLKN 204
 QY 187 LGGKLRFTYGVVDLGGGSQVMTYAVSRNTAKNAKPVPEGEDPYIKKVLQKKYDLV 246
 DB 205 LKRTTCKPYSDFGVVDLGGGSQVMAVISEEDAAAPKPLEGEDSVREMYLGRKRYFL 264
 QY 247 YHSTYLRVREARERATFEFYVAGGSANPCILAGDGAITYSGAITYKSAPASGSLNOCRI 306
 DB 265 YHSTYLRVREARERATFEFYVAGGSANPCILAGDGAITYSGAITYKSAPASGSLNOCRI 324
 QY 307 IALKALV-NAPCPYONCTFGGIMNGGGSGQKNLFTSSFYLSFVLSDEVGIFV--NPNKAK 363
 DB 325 VAINALKVNNSLCTHMKCTFGGVWNGGGGQKNMFVASFPPRAAEAG-FVDPNQPVAIE 383
 QY 364 IRPVDLKTAAKLACKTNLEDAKSKYIPDLYEKDSVEYVCLDLYVYLLVDGFGLDPPQEV 423
 DB 384 VRPLDEKKAARAKCSKMLEGKSTFP-EEDNLPYLCDLVYQYTLILDGFGLEPSQITLV 442
 QY 424 TVANEIEYODALVEAAMPGLGTAIEAISS 451
 DB 443 TLVKKRYGDAVEAAMPGLGSAIEAVSS 470

RESULT 14
 ID 09M7B3 PRELIMINARY; PRT: 472 AA.
 AC 09M7B3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE APYRASE.
 GN ATAPY2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 01:31:50 ; Search time 2851 Seconds
(without alignments)
10929.363 Million cell updates/sec

Title: US-09-657-631-8
Perfect score: 1489
Sequence: 1 aagtgctctctctctctgtag.....cgcattctctctctctt

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: GenEmbl:*
2: gb_ba:*
3: gb_htg:*
4: gb_in:*
5: gb_cm:*
6: gb_ov:*
7: gb_pat:*
8: gb_ph:*
9: gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1489	100.0	1489	8	AF156780	Lotus jap
2	889.8	59.8	1463	8	AF207688	AF207688 Glycine s
3	876.8	58.9	1401	8	AF288132	AF288132 Medicago
4	869.2	58.4	1460	8	AF156782	AF156782 Medicago
5	844.6	56.7	1648	8	AB027613	AB027613 Pisum sat
6	844.6	56.7	1651	8	AB038668	AB038668 Pisum sat
7	839.8	56.4	1572	8	AB027614	AB027614 Pisum sat
8	839.8	56.4	1645	8	AB022319	AB022319 Pisum sat
9	839.8	56.4	1648	8	AB038669	AB038669 Pisum sat
10	839.8	56.4	1661	8	PSMTPEASE	232743 P. sativum (
11	835	56.1	1667	6	ES1056	ES1056 Disease res
12	833.4	56.0	1697	6	ES1055	ES1055 Disease res
13	830.4	55.8	1368	6	ES1054	ES1054 Disease res
14	830.4	55.8	1557	8	AB027616	AB027616 Pisum sat
15	824.6	55.4	1608	8	AF139807	AF139807 Dolichos
16	782.4	52.5	1423	8	AB027615	AB027615 Pisum sat
17	665.6	44.7	1395	8	AF207687	AF207687 Glycine s
18	652.4	43.8	1435	8	AF156781	AF156781 Dolichos
19	630	42.3	1060	8	AF288133	AF288133 Medicago
20	628.2	42.2	1790	8	AF305783	AF305783 Pisum sat
21	500.4	33.6	1827	8	AF141671	AF141671 Arabidops
22	500.4	33.6	1953	8	AF156783	AF156783 Arabidops
23	470	31.6	2024	8	AF093604	AF093604 Arabidops
24	441.4	29.6	1530	8	STU58597	STU58597 Solanum tub
25	250	16.8	2216	8	AB023621	AB023621 Pisum sat
26	250	16.8	2350	8	AB038554	AB038554 Pisum sat
27	250	16.8	2350	8	AB038555	AB038555 Pisum sat
28	194.4	13.1	751	8	AB032754	AB032754 Solanum m
29	115.8	7.8	620	8	AB030444S1	AB030444 Pisum sat
30	112.2	7.5	100814	2	AP003929	AP003929 Oryza sat
31	110	7.4	584	8	AB030444S2	AB030444 Pisum sat
32	106.6	7.2	83948	8	AB012246	AB012246 Arabidops
33	97.2	6.5	1900	8	AF465240	AF465240 Schizosac
34	97.2	6.5	21917	8	SPAC824	AL131741 S. pombe c
35	97	6.5	1800	8	CAL421721	AJ421721 Candida a
36	93.8	6.3	110804	8	ATAC016829	AJ238636 Arabidops
37	93.2	6.3	1633	10	NMU238636	AJ238636 Mus muscu
38	93.2	6.3	1939	10	AF136571	AF136571 Mus muscu
39	93.2	6.3	2099	10	BC015247	BC015247 Mus muscu
40	93.2	6.3	2119	10	AF006482	AF006482 Mus muscu
41	91.2	6.1	1998	9	AF039918	AF039918 Homo sapi
42	91.2	6.1	2033	9	AF136572	AF136572 Homo sapi
43	91.2	6.1	3157	9	BC020966	BC020966 Homo sapi
44	85.6	5.7	1749	10	AF084568	AF084568 Mesocric
45	85.6	5.7	1750	10	AF084569	AF084569 Mesocric

ALIGNMENTS

RESULT	1	AF156780	1489 bp	mRNA	linear	PLN 09-NOV-1999
LOCUS	AF156780	Lotus japonicus nod factor binding lectin-nucleotide				
DEFINITION	AF156780	Lotus japonicus nod factor binding lectin-nucleotide				
ACCESSION	AF156780	phosphohydrolase (LNP) mRNA, complete cds.				
VERSION	AF156780.1	GI:6006794				
KEYWORDS						
SOURCE		Lotus japonicus.				
ORGANISM		Lotus japonicus				
REFERENCE		Lotus japonicus				
AUTHORS		Roberts,N.J., Brigham,J., Wu,B., Murphy,J.B., Volpin,H., Phillips,D.A. and Frzler,M.E.				
TITLE		A Nod factor-binding lectin is a member of a distinct class of apyrases that may be unique to the legumes				
JOURNAL		Mol. Gen. Genet. 262 (2), 261-267 (1999)				
MEDLINE		99444909				
PUBMED		10517321				
REFERENCE		2 (bases 1 to 1489)				

AUTHORS

Roberts,N.J., Brigham,J., Wu,B., Murphy,J.B., Volpin,H.,
Phillips,D.A. and Etzler,M.E.

TITLE

Direct Submission Molecular and Cellular Biology, University
Submitted (07-JUN-1999) of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA

JOURNAL

FEATURES

Location/Qualifiers
1. 1489
/organism="Lotus japonicus"
/cultivar="Gifu"
/db_xref="taxon:34305"
/class="type="root"
1. 1489
/gene="LNP"
43. 1413
/gene="LNP"
/function="apyrase"
/codon_start=1
/product="nod factor binding lectin-nucleotide
phosphohydrolase"
/protein_id="AAFO0609.1"
/db_xref="GI:6006795"

source

gene

CDS

translation="MDRLSLMTFVFMPLPAISSOYLGNINIMNKILPKNOEYV
STAVLPDAGSTRVHVNFDNDLIPVENEDEFYDVKPGLSSTYANPEEAESLI
PLKKAENVVPSQDPNTPVKLGATAGLLEGNAENILQAVRDLMSRSALNVSD
AVSILDQEGESLWVINYLLGKLEKRFKTVGVNLDGGSVOMTAVSRNTAKNAP
KYPEGEDPYIKKLVLQGGKDYLVHSHLYRGRFAEFKFKVAGSANGCIIAGFDGA
VYSGAEYKVASAPASGSLNOCRTIKALKVAPPCPNOCNTEFGGIMNGGGSGOKNL
FLTSSEYVLSDEVGTFVFNKPNKIRVPDKTAKKINLEDAKSEYDYLKESDVE
YVCLDLYVYITLIVDSFGIDPFOEYVANEIEYODALVRAANPLGTAITAISSLPFE
RLMPLFI"

BASE COUNT
ORIGIN

437 a 284 c 324 g 444 t

Query Match

100.0%; Score 1489; DB 8; Length 1489;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aagtgctctctctctgtagtagtgtagtgcactgaagcgaatgaagccttcaatagt 60
DB 1 AAGTGCTCTCTCTCTGTAAGTAGTTCGATTCGATTAACCATGACCTTCTTAATAGT 60
QY 61 ctcatgaaccttggtcatggttaatgctgctactctctccccaataatctcgaaac 120
DB 61 CTCATGACCTTTGTGTTCATGTTAATGCCGTATCTCTCCGCCAATATCTCGGAAC 120
QY 121 aaattctcatgtaatgtaagatatattctcccaaaaaatcagaagcaagttaactac 180
DB 121 AACATTCTCATGTAATGTAATGTAATCTCCAAAAATCAGAACACAGTTACATCTATC 180
QY 181 gctgtatattgtagtctgtagtaactggaagcagatgcatactgtcaatttgatcacg 240
DB 181 GCTGTATATTGATGCTGTGTAGCAGCTGGAAGCAGAGTCCATGCTTCAATTTTGATCAG 240
QY 241 aactatgatctcctcccggttgaacgaacgaactgaagtttatagtatcggttaaacccgt 300
DB 241 AACTATGATCTCTCTCCCGGTTGAAAAAGAACTTGAGTTTATGATGTGTTAAACCCGCT 300
QY 301 ttgaattcaagcgtgtaactctgaagaagcgtgcaagaatctcgtatccaatttcaaaa 360
DB 301 TTGAGTTCAATGCTGTATATCTGGAAGAGCTGCAAGAACTCTGATTCCACTTCTTAAA 360
QY 361 gaagcagaataatggtctcctgtagcgaacccaacccaaccccggttgaagcgttgagca 420
DB 361 GAAGCAGAATAATGTGTTCTGTGTAGCAGCAACCCACACCCCGTTAAGCGTGGGCA 420
QY 421 aactgcaagtttaagcgttttgagaggggaatgctgctgtaaaatatatgcaagcgctcag 480
DB 421 ACTGCAAGTTTAAGCGTTTGGAGGGGAATGCTGTAATAATATGCAAGCGCTGAG 480
QY 481 gatatctcagaagaagtgcccttaagtccaatcgaatcgaatcgaatcgaatcgaatcgaat 540
DB 481 GATATCTCTCAGCAAGAAAGTCCCTTAATGTTCATCAGATGATCTATCTTGTAT 540

QY 541 ggaacccaagaagttcttattcttgggtgacaaataactatctcttggggaagtggga 600
DB 541 GGAACCCCAAGAAGTTTATTCTTTGGGTGACAAATTAATCTCTTGGGGAAGTTGGGA 600
QY 601 aaaagatttcaagaagaagtgagtagtgatcctagaagagtggttcagtgcataatgaca 660
DB 601 AAAAGATTTCAAAAGACAGTGGGAGTAGTGTATGATCTAGAGAGTGGTCAAGCAATGACA 660
QY 661 tatgcaagttccaaagaacaaagcttaaaatgctcccaaaagtactgaagaagggatcca 720
DB 661 TATGCAAGTTTCAAGAACACAGCTTAAATGCTCCAAAAGTACTGGAAGAGAGATCCA 720
QY 721 tacataagaagcgttgaactcccaagggaagaataatgaacttattgcttcaacttctg 780
DB 721 TACATTAAGAAGCCTTGACTCCAGGGAAGAAATATGACTTTATGTTCAAGTACTTGTG 780
QY 781 cgtatggaagaagaagcattcgtgcagagatttcaagtcgctggtggttctgctaact 840
DB 781 CGCTATGGAAGAGAGCATTTGTCGAGAGATTTTCAAGGTGCTGCTGCTGCTGCTAAT 840
QY 841 ccttgcaatttgaagcgttgaagtggtggtggtggtggtggtggtggtggtggtggtg 900
DB 841 CTTGCAATTTTACCTGCTGCTTGAATGGGCAATATACATATTCGAGCAGAGATTAAGGTC 900
QY 901 tggcccccaagcctcaagatctaaacttgaatcaatgcaagaagaatagcttcaagcctct 960
DB 901 TCGGCCCCACCTTCAGAGATCTTAATCAATCAATGCAAGAAAGATAGCTTTAAGGCTCT 960
QY 961 aaagtgaatgacaccttgctccctatcagaatctgaacttgggtgggataatggaatgga 1020
DB 961 AAAGTGAATGCACCTTGTCTCTTCAAGAAATGCACTTTGTGGGATATGGAATGGTGA 1020
QY 1021 ggtggaagtggtcaaaaaaacttctctacttcaacttcttcaacttctgaagat 1080
DB 1021 GGTGGAAGTGTCAAAAAATCTTCTTACTTCACTTCTTCTTCTTCTTCTGAAAGAT 1080
QY 1081 gttggatcttctgtaaaacccaatgccaatctgctcagattgatttgaagaactgca 1140
DB 1081 GTTGGATCTTTGTGAATTAACCAATGCCAAATGTCGACAGTGTGATTAAGACTGCA 1140
QY 1141 gctaaactagcttgtaaaacaaatcttgagatgcaaaatcccaatcccaagcttctat 1200
DB 1141 GCTAAACTACTTGTAAACAAATCTTGAGAGTGCAAAATCCCAATCCAGATCTTTAT 1200
QY 1201 gagaagaagctgttgtaataatggtgcttgatcttctgaagcttctgaagctgctgtt 1260
DB 1201 GAGAAAGACAGTGTGTAATATGTGTGCTTGGATCTTGTACGTGACATTCCTTGT 1260
QY 1261 gatgatttgcttgaatccatttcaagaaggttacagtggtgcaatggaatggaatcacg 1320
DB 1261 GATGATTGTGCTGTGATTCATTCAAGAGGTTAACATGCGAATGAATGAATATCAG 1320
QY 1321 gatgctcttgtaggaagcgaatggtcctctaggaactgcaatagaagaatatatactatg 1380
DB 1321 GATGCTCTTGTGGAAGCCGATGGCCCTTAGGCACTGCATAGGAAGCAATATATATCTTGT 1380
QY 1381 cctaaatttgagagatttaagttatttatttaactaactgaactgaactgaactgaactga 1440
DB 1381 CCTAAATTGAGAGATTAAATGATTTATTTAACTGATCTGATCTGATCTGATCTGAT 1440
QY 1441 taccatgaagaataaataatgaataaagaacccgaacttcttctcctgtt 1489
DB 1441 TACCTGAAGAATAAATAATGAATAAAGCCGATCTTCTTCTCTGCTT 1489

RESULT 2

AF207688 1463 bp mRNA linear PLN 18-NOV-2000
LOCUS AF207688
DEFINITION Glycine soja apyrase GS52 mRNA, complete cds.
ACCESSION AF207688
VERSION AF207688.1 GI:11225136
KEYWORDS

SOURCE Glycine soja.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 1463)
AUTHORS Day, R.B., McAlvin, C.B., Loh, J.T., Denny, R.L., Wood, T.C., Young, N.D. and Stacey, G.
TITLE Differential expression of two soybean apyrases, one of which is an early nodulin
JOURNAL Mol. Plant Microbe Interact.. 13 (10), 1053-1070 (2000)

MEDLINE 204966172
PUBMED 11043467

REFERENCE 2 (bases 1 to 1463)
AUTHORS Day, R.B., McAlvin, C.B., Loh, J.T., Fink, S.E., Denny, R.L., Dunlap, J., Young, N.D. and Stacey, G.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-1999) Microbiology, The University of Tennessee, M409 Walters Life Science Building, Knoxville, TN 37996-0845, USA

FEATURES
source Location=Qualifiers
1..1463
/organism="Glycine soja"
/db_xref="taxon:3848"
/chromosome="MIG_J"
/lssue_type="root"
15..1421
/codon_start=-1
/product="apyrase GS52"
/protein_id="AAG32960.1"
/db_xref="GI:11225137"
/translation="MVLVMSRGKKNDFLTFTLLLFHTALSTOYHDNILLT
HRIIFPKQEAITSVAVFEDAGSTGRVHVHFNDLRLRGNLEFERYDKTPGLSAW
ADNPQQAESLIPLLEAESVVEDLYTPVPKLGATAGLRLEGDSERTLOAVRD
LKRNSTNAQPDAAVIDDGOEBSYLWTINYLIGRKRESKRVGVDDGGSSVOMA
KAISRYNNAKNAPRPDGEERYIKLVLYKGREDIYHSYHLHGEMASRAELIKATDGS
ANPEILLDGYDTYSYGADYFAFAPKIPSGSSYDECREVVLAOKINESCPHNCFFGG
WDGREGSGOKILFPGTSFYLYLPETGIIDLNKPMSKHVPDLIEARACCTKLDAKS
STYPNPAEDRLPYCLDIAYVALYTGDGFSHPQWEITVANEIYQDALVEAAPLGG
ALEAISLPKFDRLMYPI"

BASE COUNT 417 a 301 c 324 g 421 t
ORIGIN

Query Match 59.8%; Score 889.8; DB 8; Length 1463;
Best Local Similarity 79.6%; Pred. No. 9.6e-199;
Matches 1105; Conservative 0; Mismatches 272; Indels 12; Gaps 4;

Oy 32 ggacataagccatggaactcttcaattagtctcatgacctgtgtgttcattgtaalgc--- 88
|||||
Db 40 GGACTAGAACAATGTGCAGACTTCCTTAACCCTCCTCACCTCTCTACTCTTGTTATTCAATTCAA 99
|||||
Oy 89 ctgtctatctctctcccnaatatctcgaaacaanattctcatgaatcgtaaagatalaac 148
|||||
Db 100 CAGACCTCTCTTTCACCCCAATATCATGATGAGGAACATCTTACTCATCTATCGTAGAACATCT 159
|||||
Oy 149 tcaccaaaaatcagaacacagttcatacatagcgtgtataatttgatgctgtagacatty 208
|||||
Db 160 TCGCCAAA---CAGAAAGCTATCACCTCGTACCGCTGTATCTTTGATGCTGTGACACTG 216
|||||
Oy 209 gaagcagagtcacatgltcacaaatttgatcagaacttagatctctccgttgtgaaaaag 268
|||||
Db 217 GTACCCGGTGTCAATGCTTTCATTTCGACCAGAACTTGATGATCTCTGATTTGCAACAG 276
|||||
Oy 269 aacttgagtttatgatctcggtttaaccgccgtttgagttacatagcgtctaactctgaag 328
|||||
Db 277 AACTGGAATTTTATGATAGGTAAAGGTAAACCCCGGTGTGTAAGTGAATGCAATGCAATATCTCTCAAC 336
|||||
Oy 329 aaagctcagaatctctgtatccaactctaaagaagacagaanaatgtggcttcctgtagacc 388
|||||
Db 337 AAGCTGCAAGATCTCTTATCCCACTTTTAAAGAGAAAGCTGAAAGATTTGTGCTCGTAGGACC 396
|||||
Oy 389 agcaaacccaacacaccgltlaagcttgggggaacatgcacagittlaagaccttggagggga 448

Db	397	TGTAACCTACTACACCCGTTAAACCTTGGGGCGACTGCAAGCTTTAAAGCTTTTGGAAAGGG	456
Qy	449	atgcgtctgtaaaatatactgcaagcgctcaaggatactgcacgaacgaagtcctcta	508
Db	457	ATGCTTCTGAAAGATATATGCAAGCGGTTAAGGATTTGCTCAAGAAACGAAGTACCTGCA	516
Qy	509	atgttcaatcagaatgcagtatctactctgatgaaacccaagaagttctactcttggg	568
Db	517	ACCGTCAACACAGATGCAAGTAGTATATTGATGAAACCCAAAGAGGTTCTATCATAGGG	576
Qy	569	tgaacaataactctatcctctggggaatttgggaaaaagatttcaagaagcaggggaag	628
Db	577	TGCAAAATTAATTAATTATTTGGGGAAGTTGGAAAAAGTTTTCAAAGACAGTAGGAGTAG	636
Qy	629	ttgatctagaaggtgggtcagctgcaaatgacatgcaatctccaagaacacagctaaaa	688
Db	637	TTGATCTTGGAGGTGATCAGTTCATTAATGGCATATGCCATCTCAAGAAACACAGCAAAA	696
Qy	689	atgctccaaaagtacctgaaagagagatcacaataaagaagcttgactccaaggaa	748
Db	697	ATGCTCCAAACACACAGATGAGAGAACCTTACTATAAGAAAGCTTTACTCAAGGAA	756
Qy	749	agaaataatgaccttatgttcaaatgacttgcgtcatgtaagaaagaagcatctgtgag	808
Db	757	AGGAATATGACCTCTATGTTACACAGTATTTTGCATTACGTTAGAGAACATCTCGTGGC	816
Qy	809	agattttcaaggtgcgctggtgtgtctgcctaaacctgtcaatctttagctgtgttggg	868
Db	817	AGATTTTGAAGGTCTACTGATGCTCCGCTTAACCTTTGTTATTTTACTGGCTACGATGGGA	876
Qy	869	catatacatatccggagacagagatataaagctctcgggcccaagcttcaagatctaactga	928
Db	877	CTTACACATATTCAGGAGGACGACTATTAAGGCCCTTCCCTCCATTTCTGGCTCAAGTACG	936
Qy	929	atcaatgcagaaagatagctcttaaggtccttaaaagtgaatgacacctgtgctcataaga	988
Db	937	ATGATGTTAAGAGGGTGTGTTCTTCAAGCTCTCAAAATGATAGTACATGCCCCCATCAAA	996
Qy	989	attgcacttttggtaggataatgaatgagtggtgggggtggaagtggaagcgaagaaattctttcc	1048
Db	997	ATTGCAATTTGGTGGATATATGGGATGTGGAAAGGGGAAGTGACAGAAATTTCTTTTCG	1056
Qy	1049	ttaactcatcttcttaactacctctctgaaagatgttggagctcttgg---tgaataaaccca	1105
Db	1057	GCACTTGCTTCTTCTATTTATCTACCTACAGAGATTTGGTATCATTTGATCTCAATAAACCA	1116
Qy	1106	atgccaaatctgltccagttgataltgaagactgacgtctaactagctgtglaaacaaatc	1165
Db	1117	ACTCCAAATATTCACCTCGTTGATCTGGAATGGAAGCTTAAGACACTGTGTGAACAAATAT	1176
Qy	1166	tttagagtgtaaaatccaataaccagatccttaagagaagaaggttttgaatagt	1225
Db	1177	TGGAGGATGCCAAGTCCACTATTCACAAACCCCTGTAG---GATAGACTTCCATATTTGAT	1233
Qy	1226	gcttgagactctgtactacgtgtgatacaactgtctgtttagatgatttgccttgatccatttc	1285
Db	1234	GCTTGGATTTTGGATATTCACAGTATGATTTGTACACGAGATGAGATTTTGAAGCTTGATCCATGCG	1293
Qy	1286	aagaaggttcaagctgagcgaatgaatatgaatatcagaatgctctgttggaaagcgcatggc	1345
Db	1294	AAGGATTTACAGTCGCAAAATGAATATGAATATACAGATGCTCTTGTGAAAGCTGATGAGC	1353
Qy	1346	ctctaggaagctgcatcagaagaacatatcatctgacctaatttgaagatataatgtatt	1405
Db	1354	CTTTAGGCACTGCATAGAAAGCCATATGCTATGCTTAATTTGATGCAATTAATGATATTT	1413
Qy	1406	ttatctaaat 1414	
Db	1414	TTATTTAA 1422	

Db	190	ACTGGAAGCCGTGCATGTTTAACCATTTTGTATCAAAACCTAAATCTTCTCATGTGGT	245
QY	265	aacgaacttgagttatgatcgtgttaaccgggttgatctcaacgctcaatcct	324
Db	250	AAAGATGTTGAGTTTATTAATGAAGCAACGCCGGTTTGAGTCATACGCGATATATCA	309
QY	325	gaagaagctgcgaacatcttgatctcactctcaagaagcagaanaatggtctcctg	384
Db	310	GAGGAAGCTGCAAAATCTTTGATCCACTTTTGAAGCAAGAGAAAGTATGTTCCGAG	369
QY	385	agccagcaaccocaacacaccggttaagcttggggaactgcaggtttaagcctttgag	444
Db	370	GATCACCGCTCCAAAGCACCCCTATGTAGCTTGGGCAACAGACGGTTTAAAGCTTTGAT	429
QY	445	gggaatgcgcgtgaanaatbatatggaacgggttcagggatgctcgaacagaatggc	504
Db	430	GGGGATGCTTCTGAAAAAATACTGCATTCGGTAAAGGATTTGTTACCAATATGAAATACC	489
QY	505	cttaatgltcaatcagatgcagatgcatctatctcttgatggnaaccagaagttcttatct	564
Db	490	TTCAATGTTCAACCTGATGCACTGACATTTCTATTTATGATGAAACCCAAAGGTTGTTATCTG	549
QY	565	tgggtacacattaacatctctcttggggaagttggaaaagattcaagaacagtgga	624
Db	550	TGGGTACAGTTACTATGCACTTTGGGGAATTTAAGGAAAAAGTTCACAAAAACAGTGGGA	609
QY	625	gtagtgaactcagagaggttgggtcagttgcanaatgacatagcatgcctcaaggaaacagct	684
Db	610	GTAAATGAGATGTTGGAGGTGATGATCAAGTTCAAAATGGCATATGACGTGCAAAAGTATACAGCT	669
QY	685	aaaaatgctcccaaaagtaacctggaagagagagatccatcacgtaagaagcttgcacag	744
Db	670	AAAAATGCTCCAAAGTGTGCTGATGAGAGACATCCATACATTAAGAAAGCTTGATCAACG	729
QY	745	ggaagaatatbgaacttattgltcaacagttacttgcgcatalggaagagagaagcttctg	804
Db	730	GGAAAAACATATGATCTCTATGTTACACAGTACTTACACTTGGTGAAGACATCTGCA	789
QY	805	ggaagagattccaagtcgcgtgcgtgttctgcgtcaactcgtcaattagctgttgt	864
Db	790	GCTGAGATTATGAAGGTCACTCGTAGTTCCTTCAACCTCTTGCCCTTTAGCTGGATTGGAT	849
QY	865	ggggcatatacatctccgagcagagtalaaagttctcgcccccagctccaagatctaac	924
Db	850	GGGACCTCAACATATGCTGGAGAGAGATTAAAGCCAAACCCCTCTCTCGGGGCTAAT	909
QY	925	ttagaactaagcagaagaataagcctctaagagcctctaagtagaactgcactgttcctat	984
Db	910	TTTATGATGATCAAAAAGATATATTCGTAAAGCACTTAAATGTGAATATCAATCAATGCTCAT	969
QY	985	cagaatgcacttltggttggagatattggaatgctgcaggttggaaatggttgcanaaaatctt	1044
Db	970	CAAAATTCACATTTTGGTGGATTTTGGAAATGGCGGAGGGGAAATATGACAGAAACACTT	1022
QY	1045	ttccctattcaatcttcctctatctcctctcctgaagatggttggagatcttgg---tgaataa	1101
Db	1030	TTTGCTCTTCTATCTTTCTTTTACCTTAACCTGAAAGATGTTGGTATGTTGATATCCAAAGACA	1089
QY	1102	cccaatgccaanaatctgltccagtttgaatttgaagactgcagctcaaaactgctgttaaaaca	1161
Db	1090	CCCTAATTTCAAAATTCGCTCGTGGATGCTGATCGTGAAGTAAAGAGCTATGTCATTTA	1149
QY	1162	aactcttgagagatgcanaaatcccaatcccaagatcttlatlgagaaagcagttgttgaatat	1221
Db	1150	AACTTTATGAGATGCAAAATCCACTTTCATCTTCATCTTCTGTAAGAAAAAATATATAGCTCATAT	1209
QY	1222	gtgtgcttgatctctgctcagctgcgtcaacatgctgttgaatgagatltgtgtcctgataca	1281
Db	1210	GTATGCAATGATCTTATATATACGATATGTGTGTTGCTGTTGATGGAATTTGGCTTATGATATCA	1269
QY	1282	tttcaagaagttacagttggcgaatgaaatlgatatcatcagagatgctctgttggaaacgcga	1341
Db	1270	TTTGCAACAAATTTCAATCGGGGAGGAATTTGAATTCGAAGATGCTGTTTGGAAAGCTGCA	1329

QY 1342 tggccttagcagctgcataagaacatcatcgtcctaattgagagattaaatg 1401
 DB 1330 TGCCCTTAGGCAATGCTGTAGAGCAATATCATCTTACCTAAATTTGACGAATGATG 1389

QY 1402 taattatttaa 1413
 DB 1390 TATTGTGTTAA 1401

RESULT 4
 AF156782 1460 bp mRNA linear PLN 09-NOV-1999
 LOCUS Medicago sativa nod factor binding lectin-nucleotide
 DEFINITION Medicago sativa (LNP) mRNA, complete cds.
 ACCESSION AF156782
 VERSION AF156782.1 GI:6006798
 KEYWORDS
 SOURCE Medicago sativa.
 ORGANISM Medicago sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 REFERENCE 1 (bases 1 to 1460)
 AUTHORS Roberts,N.J., Brigham,J., Wu,B., Murphy,J.B., Volpin,H.,
 Phillips,D.A. and Etzler,M.E.
 TITLE A nod factor-binding lectin is a member of a distinct class of
 JOURNAL apyrases that may be unique to the legumes
 MEDLINE Mol. Gen. Genet. 262 (2), 261-267 (1999)
 PUBMED 99444909
 REFERENCE 2 (bases 1 to 1460)
 AUTHORS Roberts,N.J., Brigham,J., Wu,B., Murphy,J.B., Volpin,H.,
 Phillips,D.A. and Etzler,M.E.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-1999) Molecular and Cellular Biology, University
 of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA

FEATURES
 source
 1. 1460
 /location="Qualifiers
 /organism="Medicago sativa"
 /db_xref="taxon:3879"
 /tissue_type="root"
 1. 1460
 /gene="LNP"
 15. 1382
 /gene="LNP"
 /function="apyrase"
 /codon_start=1
 /product="nod factor binding lectin-nucleotide
 phosphohydrolase"
 /protein_id="AAF0061.1"
 /db_xref="GI:6006799"
 /translation="MEFLITLITATLFLLLMPATISSOYLGNLLTRNKIFOKOELTTS
 LAAVFDAGSTGVHVFHFDNDLHLHNGDIEFVDAIKPELSAGDPEAAKSLIP
 LLEAEADVPEDLHPKTRPLRGATAGLILKNGDAEAKILQATRNHFSRSLNLRDA
 VSIIDGQEGSVMTVNVYVIGNAGSTKSTGVTIDGGSSVOMTAVNSKRTARNAP
 VADGEDPIKILVYAKGOYDYHVSILRFEGEATRAOYLANITNSANPCILPENGTF
 TYSVEYRKARSPSSGNSNDCKEILILKLVKNDPCPYSCTFGGIWNCGSGQKLF
 VTSAFVLAEDVGVNEPKPNSILHPVDFEILKARACALNEDVAKSYPRILDAKRPY
 VCMDLVQHVHLVHGFGLGPRKEITVCEGLOYONSVEAAMPCLATAVAISALPKFKR
 LMYFI"

BASE COUNT 440 a 268 c 308 g 444 t
 ORIGIN

Query Match 58.4%; Score 869.2; DB 8; Length 1460;
 Best Local Similarity 77.2%; Pred. No. 6.6e-194;
 Matches 1112; Conservative 0; Mismatches 313; Indels 15; Gaps 4;

QY 34 actaaagcagtgactcttaattagctcat---gaccttggtgctcatgttaagcct 90
 DB 6 ATTAAAGAACATGAGCTTCTTAATACATCATTCGCACTTTTCTTACTCTTGTAAATGCCT 65

QY 91 gctatctcttcctcccaatatactcggaaacaacattcctatgaatcgttaagattatctc 150
 DB 66 GCATATCACTTCTCTCCCAATATTTAGGAACACCTTCTCATATGAGAAATTTT---- 121

QY 151 cccaataatcaggaacagctacatcatacgcgtgtataattgagtcgtgtagcactgga 210
 DB 122 --CCAAACCAAGAAACCTTAACCTCTTACGCTGATATGATGATGATGATGATGATGAT 179

QY 211 agcagaagtcacagtcacaaatttgatcagaacttagatctccctccgttgaacgaa 270
 DB 180 ACTCGGTGTCAGATTTTACCATTTTGTATGATCAAGCTTGTGATCACTTGTGCAATGAT 239

QY 271 ctgagtttatgattcgttgtaacccggttgagttcagttcagttcagttcagttcagtt 330
 DB 240 ATTGAGTTTGTGACAAAGATCAACACGAGTTTGTAGTGATGAGGATATATCTGACAA 299

QY 331 gctgcagaaatctctgatactcactctcaaaagaagcagaatggtgtctcgtgagccag 390
 DB 300 GCAGCAAAATCT 359

QY 391 caaccacaacacccglttaagcttgaggcaactgcaagtttaagcttggaggggaat 450
 DB 360 CACCCCAAAACACCCCTTAGGCTTGCGGCAACCGCAGTTTGAGCTTTGGAATGGGAG 419

QY 451 gctgctgaaatataatttgcaagcgtgcaaggtatgctgaagcaacgaagtgccctaat 510
 DB 420 GCTGCTGAAATATATTTGACGACGACAAAGATATTTGACGACGAGATGACCTCAAC 479

QY 511 gttcaactcagatgcaactatctcttgatggaacccaagaaggtctctatctcttggtg 570
 DB 480 GTTCAACGATGATGCAATTTCTATTTATTTGATGGAACCCAGAAAGGTTCTTATATGTCGG 539

QY 571 acaattactatctcttgagggaagttgggaagaattacaagaacagtgaggagtagt 630
 DB 540 ACAGTTAACTATGATTTGGGGAATTTGGGAAACGTTACAAACACAGTGGGAGTAAT 599

QY 631 gctcagagagtggtggtcagatgcaaatgacatgcaagtcagtcagtcagtcagtcag 690
 DB 600 GACCTTGGAGTGCTTCACTTCAATGACATATGACATGACATGACATGACATGACAT 659

QY 691 gttccaaagaactcgtgaagaagagatccatcataaagaagcttgtaactccagggaaag 750
 DB 660 GCTCCTAAAGTGTGATGATGAGAGGATCCATATATTAAGAGCTTGTGTCAGAGGAAG 719

QY 751 aaatagaccttaagtgctcaagtaacttgctgctatggaagaagaagcaatctgctcagag 810
 DB 720 CATATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779

QY 811 atttcaagtgctggtggtggtctgctgtaactccttgatcttgatcttgatcttgatctg 870
 DB 780 GTTTGAAATGCAACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839

QY 871 tatacatattcggagcagagatataaagtgctcgtgccccagcttcagagactaactgaat 930
 DB 840 TTTCATATTTACAGAGATGAGATATTAAGGCTTTTTCCTCTCTCTCTCTCTCTCTCT 899

QY 931 caatgcagaagaatgctcttaagctcttaagtgaaatgcaacctgtgtccctatacgaat 990
 DB 900 GATTGCAAAAGATATATTTCTTAAGGTTCTTAAGATTAAGATTAAGATTAAGATTA 959

QY 991 tgcacttctgtggtggaataagaaatggtggaagtggaagtggaagtggaagtggaagtg 1050
 DB 960 TGCACCTTTTGTGGAATATGAAATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1019

QY 1051 actcatcttcttactactcctcctgaagaatggtggaatcttggtg---ataaaccat 1107
 DB 1020 ACTTACGCTTGTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079

QY 1108 gccaaaatctgctcagtgatcttgaaagctgcaagctcaactgctgttaaaacaactct 1167
 DB 1080 TCCATACCTTCATCCAGATGATATTCGAATTTGAAGATTAAGAGCTTGTGCTATTAACCT 1139

OY	810	gatttcaagtcgcgtctggatcgtccaaacctcgactttaacgttgctttgaaggagc	869
Db	793	GATTTTGAAGCTCATCCTCCGTCGTTCCTTAACCCCTTGCCCTTTAGCTGGATTAAAGAAT	852
OY	870	atacatactlccggagcagaatlataagcttcgcggcccaccgtccaagtcatcaactga	929
Db	853	CATAACATATTCAGGAGAAAGACTTTAAGGCACACTGCTTACACTTCTGGGCCAACCTTAA	912
OY	930	tcaatgccaaaagatatgctcttaagctcttaaglttaaigtacaactgtccctacaaga	989
Db	913	TAAAGCAAAAAACAAATTCGTAAAGCGCTTAAGTTGAACTATCCCTGTCCATACAGAA	972
OY	990	ttgcacttttgtgatataigatcgyrtggsgybtggaagtygtcaaaaaacttlctct	1049
Db	973	JTTCACCTTTTGGTAGAATTTGGAAATGGTGGAGAGAAATGGACAAAACCTTTTGCC	1032
OY	1050	factcaccttctctaatacctctcctgaagatgttggatctc---tgtgaataaaccaa	1106
Db	1033	JTTCCTCACCTTTCTTTTACCTFACCTGAAGATACC GGATMGTTTTGATGCAAGCACACTAA	1092
OY	1107	tgcacaalltgcctcagttgatttgaagactgcgcgtacaactagctgtttaacaactc	1166
Db	1093	TTTCACTCTTGCGCGCGGTGATATTGAACACTAAACCTTAAGAACCTTCGCGTTAACTT	1152
OY	1167	tgaagatcaaatcaaccaataccagactccttatgagaagaacagtgftgaatagtgtg	1226
Db	1153	CGAGATGCGGAANCTACTTATCCATTTCTTGATATAAAAAATGACCTTCATATGATG	1212
OY	1227	cttgaactctgtcactgtagtacatctgccttggttgatgtgattgtgtcttgatccattca	1286
Db	1213	CATGATCTTTATATATACGTATGTGTATCFGTGATGGATTTGGCTTGATCCATTTGCA	1272
OY	1287	aagagttcaagtgysgaatlgaatgaataacagatgctcttfgysgaagccgsatgsc	1346
Db	1273	AAGATTTACATCAGGAAGGAATTTGAATPACCAAGATCTATTGTGGAAAGCTGATGSC	1332
OY	1347	ctcagacactgccatagagaatalcalactlgctctaatttgagagatatgtattt	1406
Db	1333	JCTAGCAATGCTGTAGAAGCATATACAGCTTACTTAATTTGAGCGATTGATATTT	1392
OY	1407	tattaactact 1419 	
Db	1393	JGTTTTAGCTTCT 1405	
RESULT 6			
ABO38668	LOCUS	ABO38668	1651 bp mRNA linear PLN 20-MAR-2001
DEFINITION	Pisum sativum mRNA for apyrase, complete cds, clone:pkS17-2.		
VERSION	ABO38668		
KEYWORDS	ABO38668.1 GI:11596080		
SOURCE	aprase.		
ORGANISM	Pisum sativum (strain:Alaska) Seeding Dark grown stem internode CDNA to mRNA, clone_1jb:pea Internode cDNA library clone:pkS17-2. Pisum sativum		
REFERENCE	1 (sites)		
AUTHORS	Shibata,K., Abe,S. and Davies,E.		
TITLE	Structure of the coding region and mRNA variants of the apyrase from Pisum Sativum		
JOURNAL	Acta Physiol. Plant. 20, 3-13 (2001) 2 (bases 1 to 1651)		
AUTHORS	Shibata,K. and Abe,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-FEB-2000) Shunosuke Abe, Ehime University, College of Agriculture; Tsurumi 3-5-7, Matsuyama, Ehime 790-8566, Japan (E-mail:abe@ehimch.agr.ehime-u.ac.jp, URL:http://web-mch.agr.ehime-u.ac.jp, Tel.:+81-89-946-9853, Fax:+81-89-946-9853)		

FEATURES	source	Location/Qualifiers
		1..1651
		/organism="Pisum sativum"
		/strain="Alaska"
		/db_xref="taxon:3888"
		/clone="PKS17-2"
		/tissue_type="Dark grown stem internode"
		/clone_lib="pea internode cDNA library"
		/dev_stage="Seedling"
		32..1399
CDS		/standard_name="Atp diphosphohydrolase"
		/EC_number="3.6.1.5"
		/codon_start=1
		/product="apyrase"
		/protein_id="BAB18895.1"
		/db_xref="GI:11596081"
		/translation="MEFLIKLITPLLSMPAITSQYIGNLLTSRKIFLKOEISST AVEPAGSGSIHVIHYHNMJDLHIGVEYVKITPGLSSTYANNEQAKSLIPI LEADQVDVDDLOPKTPVRIGATAGRLRLNDASEKILGSVRDMISNSTPNVQDIA SIIDTOEGSYLWMTVNALNGLKRYKTVGVIDLGGSVQMAVAASKAKAPAKP ADGDDYIKKTVLGIPIPYDLVSHYSLHGRASRAEILKIPRSPNCLFAGNGITYT YSGEERKATAYTSGANENKFNILBKALKVPCYONCTFGGIIMNGGOKMLFPE SSSFPELPEDTGMNDASTPNFLRPVDETLEKAKVCAINFEPASTYFFLDDKNVAS VOMDIIQYVLVLDGFDLPLOKLTISGEKEIDAEALPEANVAISALPFEF LMFEV"
		1434..1439
BASE COUNT	497 a	311 c 308 g 535 t
ORIGIN		
Query Match	56.7%	Score 844.6; DB 8; Length 1651;
Best Local Similarity	76.5%	Pred. No. 4e-188;
Matches 1065; Conservative	0; Mismatches 319; Indels	9; Gaps 2;
30	ttggactaaagcagatgacctctcttaattgctcctatgaccttgcttgcttaatgac	89
19	ttaccctacacaaatgcagacttcccttaattaaactatnactctttctactatttctatgcc	78
90	tgcatactctccctcccaatcctcgaaacaacattctcatgaaatcgtaatatctact	149
79	tgcatacacttctcccaatcttcccaatcttgcgaacacactacacacagatgaaatcttct	138
150	ccccaaaatcaggaaacccgttaacatcagaacgttataattgattgctcgtgaaacattg	209
139	-----AAACAAGAGGAATTTCCCTTACGCTGCGATTGCGATGCTGTGAGACCTGG	192
210	aagcagatgccatcgtcacaatttgcatacgaactatagatcctccctccgttgaacaaga	269
193	tatgtccattcaatgatttaccatttttaaccagaaattagattcttctcatatgttgcaaaagg	252
270	acttgagtttatatgattcggttaaacccggtttgagttcatacgcctgctaactccgaaga	329
253	tctgcggttttataaataatataacaccttggttgagttcatgcgttaataatccagaaaca	312
330	agctgcagatcttcgatctcaactctcctaagaagaacagaanaatlggttccgttgagcca	389
313	ggctgcaaaatctctatctccatttccattttagacagacagaaatgctgctcccgacagattct	372
390	gcaaaccaacacaccccgtaagcttggggcaactgcgaagtttaaggctttggaggaggaa	449
373	tcaaccacagacacccgttagacttggggcaactgcgcggttttaaggcttttaaggaggaa	432
450	tctcgtcgaataatataatcagaagcggtcaggagatctgctcacacagaagcgcccttaa	509
433	tgccttctgaaaaatatttccaaatgcgaatggatattgctgacacacagaaatgacttca	492
510	tgttcaatagaatgcatactcattcttgaatgaacccaagaagttcttaactcttgggt	569
493	cgttcaaacacagacgagcttttataatgatggaaccccaagaaggtttatctatmggt	552
570	gacaaatcaactctctcttgggggaagtttgggaaaaagaattcaaaaaagacgttggagtgct	629
553	gacagtttaacttgcatttgggaattttgagaaaaaagatcacaaaaaacggttggagttat	612


```

Db 190 TAGTGGCATTCATGTTTACCAATTTTAAACCAACCTTAATCTTTCTTCAATTAATGGCAAGG 249
Qy 270 acttgaattatagatcttggttaaaacccggtttagtgcatacgcgtgaatcctgaaga 329
Db 250 TGTGAGATTAATTAATAGATTAACACCTGTTGAGTTTCATACGCTTAATATCCAGAAACA 309
Qy 330 agctcagaatctcgtgattccacttctaagaagcagaataatgtgtcctgtgagcca 389
Db 310 GCGTCGCAAAATCTCTCATTCACCTTTTAGAGCAAGCAAGATGATGTCGCCCGGAGATCT 369
Qy 390 gcaacccacacacccggttaagcttgggcaactgcagctgaagcttgggagggaa 449
Db 370 TCAACCCAGACACCCGTTAGACTTGGGGCAACGCGGTTTAAGGCTTTGAATGGAGA 429
Qy 450 tgcgtcgtgaataatatttgcgaagcggtcgaagatatctcagcaacagaagtgcccttaa 509
Db 430 TGCCTTGTAAAGATTAATGCAATCGTAAGGATATGCTGAGCAACAGAAAGTACCTTCAA 489
Qy 510 tgttcaatcagatgcagatctatctcttgatgtagaccagaagaagcttcttcttggt 569
Db 490 CGTTCACCAACGACGCACTTTCTTAATATGATGGAACCCAGAGAGGTTCTTAATCTATG 549
Qy 570 gacaattaactatctcttggggaagttgggaaaaaatttacaagaacagtgaggatgt 629
Db 550 GACAGTTAACTATGCAATTTGGGAATTTAGGGAAAGTACACAAAGACAGTTGAGTAA 609
Qy 630 tgcctcgaagctggtgcagtgcaaatgacatctgcagcttcaagaagcaacagctaaaa 689
Db 610 AGATCTTGGAAGTGGATGATCACTTCAAAATGGGTAATGCAAGTAAAGAACTGCTAAAA 669
Qy 690 tgcctcacaagaactgaaggaagagaatccatacaaaagaagcttactccagggaaa 749
Db 670 TGCCTCAAAAGTTGCAATGAGATGATCCATCATCAATCAAGAGTTCTACTCAAGGAA 729
Qy 750 gaaatagaccttattgttcaacagttacttgcgtatgtagaagaagcalttcgtgaga 809
Db 730 ACCATATGATCTATGTTTACAGTTACTTCACTTGGTAGAGAAAGCATCTGAGACAGA 789
Qy 810 gatttcaagctcgtgtgttctgttaatccttgcaatttgcgttgcgttgcgttgcgttgcgt 869
Db 790 GATTTGAACCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 849
Qy 870 atatacatcttcggagcagaagataaagtcctgcggcccaagcttcagatcacttga 929
Db 850 CTATACATTAATGAGAGAAAGTTAAGGCACTTACACTTGTGTCGCAAACTTAA 909
Qy 930 tcaatgagaagaagatgccttgaagcttgaagaagatgcacaccttgcctcattcaaga 989
Db 910 TAAATGCAAAACCAATTCGTAAAGCTTTAAAGTTGAAGTAACTATCTGTCATATCAGAA 969
Qy 990 ttgcacttcttgggagatagaaatgttggaggttggaggttggcaaaaaaatcttccct 1049
Db 970 TTGCACCTTTGTTGGAATTTGGAATGTTGAGAGAGAAATGAGCAAAAACCTTTTTCG 1029
Qy 1050 tacttactcttcttactcctcctgaagaatgttggatcct---tgtataaacccaa 1106
Db 1030 TTTCTATCTTCTTCTTACCTCACTTAAGATACCGGATGTTGATGCAAGCACACCTTAA 1089
Qy 1107 tgcacaatctcgtccagcttgaattgagaagctgcagtaactagcttgaataaacaatct 1166
Db 1090 TTTCACTACTTGGCGCGGTGATTTAGACTTAAGCTTAAGAAAGCTGGCGCTTAACCTT 1149
Qy 1167 tgaagtgcaaaaatcccaaatccagaactcttatgagaagaacagtggtgaatgtgtg 1226
Db 1150 CGAGATGCGAAATCTACTTATCCATTTCTTGATAGAAATAATGATGCTCATATATATAG 1209
Qy 1227 ctggaatctgtctagctgtacacatgtctgttgaatgtgtgtcttgatccattca 1286
Db 1210 CAGGATCTTAATATCACTATGATGCTTACTGATGATTTGGCTTGAATCCATTTTGA 1269
Qy 1287 aagagtgtaacgtggcgaagtgaatgaatcagatgctcttgggaagccgcatgtgccc 1346

```

```

Db 1270 AAACATTACATCAGGGAAGGAAATGCAATACCAAGATGCTATTGTGGAAGCTGCATGGCC 1329
Qy 1347 tctaagcagcgcataagaaacaatcatcatcattcgttaattgagaattgaatgtatt 1406
Db 1330 TCTAGGCAATGCTGTAGAACCCATATCAGCTTACCTTAATTTAGAGCATGATGATTTT 1389
Qy 1407 tatttaactact 1419
Db 1390 TGTTTAAGTTTCT 1402

RESULT 9
AB038669 1648 bp mRNA linear PLN 20-MAR-2001
LOCUS AB038669
DEFINITION Pisum sativum mRNA for apyrase, complete cds, clone:PKS30-3.
ACCESSION AB038669
VERSION AB038669.1 GI:11596082
KEYWORDS apyrase.
SOURCE Pisum sativum (Strain:Alaska) Seedling Dark grown stem internode
cDNA to mRNA, clone:lib:pea Internode cDNA library clone:PKS30-3.
ORGANISM Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.
REFERENCE 1 (sites)
AUTHORS Shibata,K., Abe,S. and Davies,E.
TITLE Structure of the coding region and mRNA variants of the apyrase
from Pisum Sativum
JOURNAL Acta Physiol. Plant. 20, 3-13 (2001)
AUTHORS 2 (bases 1 to 1648)
Shibata,K. and Abe,S.
DIRECT SUBMISSION
SUBMITTED (22-FEB-2000) Shunnosuke Abe, Ehime University, College
of Agriculture; Tarumi 3-5-7, Matsuyama, Ehime 790-8566, Japan
(E-mail:abe@mb.agr.ehime-u.ac.jp,
URL:http://web-mcb.agr.ehime-u.ac.jp, Tel:+81-89-946-9853,
Fax:+81-89-946-9853)

FEATURES
source Location/Qualifiers
1..1648
/oranism="Pisum sativum"
/strain="Alaska"
/db_xref="taxon:3888"
/clone="PKS30-3"
/tissue_type="dark grown stem internode"
/clone_lib="pea internode cDNA library"
/dev_stage="Seedling"
29..1396
/standard_name="ATP diphosphohydrolase"
/EC_number="3.6.1.5"
/codon_start=1
/protein_id="BAB18896.1"
/db_xref="GI:11596083"
/translation="MEDLIKLIFLFSMPAITSQYLGNLLTSRKIFLKOEISY
AVVDAGSTGSRIVHYHFNLDLHIGKEVEYNNKTPGLSSYANPBOAASLPL
LEOAEVVPDLOKPTPVRLGATGAGTGLNGDASEKILQSVRMLSRSPFNOPAV
SIIDGQEGSLIWTVAIGNAGKKTGTVGVYIDGGGSVQAAVAVSKTANABKV
ADGDDPYIKRVVYKGIYDLYVHSYHFGESABAEIKLTPSPNCLAGNGIYT
YSGEFAATATYSGAANPKCKNTIRKALNTIPCPYONCTFGGIMWGSGGNKLFY
SSSFYLPEDTGWDASTPNFILLRPVIEFKAEACALNEDAKSTYPPFLDKKNVAVS
VCMDLIYOYVLLVDFGLDPLDKITGSEKIEYODAIVEAMPIGNAVEAISALPKFER
LMYFV"
BASE COUNT 497 a 312 c 308 g 531 t
ORIGIN

Query Match 56.4%; Score 839.8; DB 8; Length 1648;
Best Local Similarity 76.2%; Pred. No. 5,3e-187;
Matches 1062; Conservative 0; Mismatches 322; Indels 9; Gaps 2;
Qy 30 ttgactaaagcagacttcttaattagctcagacttctgttcaatgtatgcc 89

```

```

Db      16 TTACCCCTACACAAATGGAGCTCCTTATTAACCTTATCACTTTCTACTATTTCTATGCGC 75
Oy      90 Tgatactctctccccaatcctcgaagacaacatctcagaatcgaatctact 149
Db      76 TGCATATCACCTCTCCCAATACTTAGAANAACACTACTCACAGTAGAANAATTTCT 135
Oy      150 ccccaaaaatcgaagaccgltacatcactcgtctatatttgatctgctgtagacatg 209
Db      136 -----AAACAAGAAGAAATTTCTCTTACGCTGATGATTCGATCGGTAGACACCG 189
Oy      210 aagcagagtcactcctacaaatttgatcagaacttagatctcctccgttgaaacga 269
Db      190 TAGGTGCATTTCAATGTTTACCATTTTACACAGAACCTAGATCTTCTCATATTTGGCAAAAG 249
Oy      270 acttgattttagtctggttaaacccggttgatgatacagctgctgaacccgaaga 329
Db      250 TGTTAGATTTTAATTAAGATTAACACCTGCTTGAATTCATACGCTAATTAATCCAGAAC 309
Oy      330 agctcagaatctcgtatctcactctcctaaagaagcagaanaatgctgctcgtagcca 389
Db      310 GGCTGCAAAATCTCTCATTTCCACTTTTAGAGCAAGCAGAGATGTCGTCGCCGAGATCT 369
Oy      390 gaaacccaacacaccccgtaagcttgaggcaactgcaagtttaagccttttgaggaggaa 449
Db      370 TCAACCCCAAGACACCCGTTAGACTTGGGGCACTCCGGTTTAAAGCTTTGAATGGAGA 429
Oy      450 tgcctcgaataatataatgcaagcgctcagggatagctcgaacacagaagtcgccctaa 509
Db      430 TGCCTCTCAAAAAGATTTGCAATCGTAAAGGATATGCGTAAAGCAAGAAAGTCTTCA 489
Oy      510 tctcaatcagaatgcaatctatctctgataatggaacccaagaagttctatcttggt 569
Db      490 CGTTAAACCAAGACGAGTTTCTATATGATGGAACCAAGAGGTTCTTATCTATGGCT 549
Oy      570 gaaataactatctctcgtgagggaagctggaagaaatctcaaacagacgtggaagtagt 629
Db      550 GACATTTAACTATGATTTGGGAAATTTAGGAAATAAGTACAAATAACAGTTGGAGTAT 609
Oy      630 tgatctaggaggctggtgtagcgaatgacataatgcaatcgaagcctcaagacaagctaa 689
Db      610 AGATCTTGAGAGTGATAGTTCAATGCGATGCGATGATGCAATGCAATGCAATGCAATG 669
Oy      690 tgcctcaaaagactcgtgaggaagagatccatcacaataaagaagcttgatctcgaagaa 749
Db      670 TGCCTCAAAAGTTGCAAGATGGAGATGATCCATCAATCAAGAGTTGATCTCAAGGAAT 729
Oy      750 gaaataagaccttcatgctcaagttactgctgctcgtcgaagagaagacatctcgtcaga 809
Db      730 ACCATATGATCTCTATGTTTCAAGTTACTTACACTTCGTTAGAGAGATCTCGAGCAGA 789
Oy      810 gatttcaaggtcgtcgtggtcgtcgaatccttgcaatcttgatctgctgcttgaggggc 869
Db      790 GATTTGGAAGTCACTCTCGTTCCTCAACCTTCCCTTACCTGATTTAAATGGAAT 849
Oy      870 atatacatctccggaagagataaggtctcggcccccagcttaagatctaaacttgaa 929
Db      850 CTATACATATTCAGAGAGAGAGTTTAAAGCAACCTTACATCTTGTGGCAAACTTTTA 909
Oy      930 tcaatgagaaagatagctcttaaggtcctttaaagtgatgcaactctccctacagaa 989
Db      910 TAAATGCAAAAACAAATTCGTAAGGCTCTTAAGTTGAACATCTCTTGTCCATATCAGAA 969
Oy      990 ttgcaacttggctggagataatggaatgctggaagctggaagctggccaataatcttccct 1049
Db      970 TTGCACTTTTGGTGAATTTGGAATGCTGAGAGAGAAATGAGCAAAAACCTTTTTCG 1029
Oy      1050 tacttcatcttctatacctctcgaagatgttggaatctt--tgtaataaaccocaa 1106
Db      1030 TTCTTCACTTTCTTTTACCTACCTGGAAGATACCGGTATGGTTGATGCAACACCTTA 1089
Oy      1107 tgcgaataatcgtcgaagttagttagaagactgcaactaactagctctgtaaaacaatct 1166

```

```

Db      1090 TTTCACTACTTCGCGCGGTGATATTGAGACTAAAGCTTAAGAAAGCTTGCGGCTTAAACTT 1149
Oy      1167 tgaagatgcaaaaatccaaataccagatcttcttaagagaagcagtggtgaatgtagt 1226
Db      1150 CAGGATGCGGAATCTACTATATTCATTTCTTANTAGAAAAATGATGCTTCAATGATAG 1209
Oy      1227 ctgagatctgtctcgtgtacacatctgcttltgatagtatctgtcgtacatctta 1286
Db      1210 CATGATCTTATATATCATATGATGTGTACTCGTTGATGATTTGGTCTGATCATGCA 1269
Oy      1287 agaagttacagtgcgcaatgaaatgtaataatcagaatgctctgtggaagcgcatgccc 1346
Db      1270 AAAGATTACATCAGGAGGAAGAAATGCAATCCAGATGCTATGTGGAAGCTGCAAGGCC 1329
Oy      1347 tctagcactgccaatagaagaacaaatcaccatctgctcaatcttgaagataatgattt 1406
Db      1330 TCTAGGCAATGCTGTGAGAACCATATACCTTACCTAATTTGAGCGATGATGATTT 1389
Oy      1407 tatttaactact 1419
Db      1390 TGTTTAAGTTTCT 1402

RESULT 10
PSNPASE
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pisum sativum
Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.
1 (bases 1 to 1661)
Hsieh, H.L., Tong, C.G., Thomas, C., and Roux, S.J.
Light-modulated abundance of an mRNA encoding a
calmodulin-regulated, chromatin-associated NTPase in pea
Plant Mol. Biol. 30 (1), 135-147 (1996)
96197404
2 (bases 1 to 455)
Hsieh, H., and Roux, S.J.
Cloning of the cDNA and expression of the mRNA for a gene encoding
a calmodulin-regulated nucleoside triphosphatase associated with
the envelope of pea nuclei
Unpublished
3 (bases 1 to 1661)
Direct Submission
Submitted (21-Apr-1994) Hsieh, H., University of Texas at Austin,
Botany, Biological Lab. Rm 6, Austin, Texas, USA, 78713
location/Qualifiers
1..1661
/organism="Pisum sativum"
/strain="Alaska"
/db_xref="taxon:3888"
/clone="p480-1-9"
/tissue_type="Plumule"
/clone_lib="Pea plumule cDNA library"
/dev_stage="Seedling"
45..1412
/gene="NTPase"
/gene="NTPase"
45..1412
/function="might be involved in RNA transport out of
nuclei"
/citation=[1]
/codon_start=1
/product="nucleoside triphosphatase"
/protein_id="CAA83655.1"
/db_xref="GI:563612"

```

```

/db_xref="SWISS-PROT:P52914"
/translation="MEDLLITFLFLFSEVYKINNTLSKRIEFLKOEISSY
AVVEDAGSTSRHIVHFNONDLHLGKVEYKYNITPELSSYANNPRAKSLIPL
LEOAEDEVPPDLOPKFPRVLRGATAGLRLNDASEITIOSRMLSNREFFNPQPAY
SLIDQDEGSYLPVWYVYALCNGKRYTKRYVYDLDGGSVOKAVISKRTANAKRY
ADDDDPYIKRVVYKGIPIYDIYVSHFGEASRAELIKTPSPNDCILAGFNGLYT
YSGEERKATVTSIGANPNKCNTRKRLKINYPYONCFGGIMWGSGNGKINFA
SSSEFYLPEDTGVNDASTPNFLLRPYDIETKAEACALNEFDASTYEPFLDKKNVASY
VCMDLIYOYVLVDFGFDLPLOKITSKEIEYODAIYEAMWPLGNVEAISLPKFER
LMYEV"
polya_signal 1447..1452
/citation=[1]
BASE COUNT 501 a 316 c 310 g 534 t
ORIGIN

```

```

Query Match 56.4% Score 839.8; DB 8; Length 1661;
Best Local Similarity 76.2%; Pred. No.5.3e-187;
Matches 1062; Conservative 0; Mismatches 322; Indels 9; Gaps 2;

QY 30 ttgactaagcgaactcttaattagtcgaatgacctgtgtgtcatgttaatgac 89
DB 32 TTACCTTACACAAATGGAGCTCTTATTAACTTATCATCTTCTACTATTCTATGCC 91
QY 90 tgcattctctcccaatatctcgaaacaacatctcatgaatcgtaagataact 149
DB 92 TGCATACCTCTCCCAATCTTAGAACAACCTTACTACCACTAGAAAGATTTCCT 151
QY 150 ccccaaatcaggaacagttacatatacgcgtgtatatttgatgtgtgtagcattg 209
DB 152 -----AAACAAGAGGAAATTTCTTACGGCTGTAGTATGATGATCTGTGACACCG 205
QY 210 aagcagagtcacatgctacaaatttgacgaactagatctctccgtgtgaacga 269
DB 206 TAGCGCATTCATGTTTACATTTTAAACCAACCTAGATCTTCTCATATTTGGCAAAG 265
QY 270 acttgattatgattcggttaacccggttggaatttcaacgcgttaacttaaga 329
DB 266 TGTGATATATTATATATAGATACACCTGTTGAGTTCACTATATATATCCAGACA 325
QY 330 agctgcgaactctctgattccacttcaaaagacgaagaatgtgttctctgtgaca 389
DB 326 GGGTCGCAAAATCTCTCATCTTCCATTTAGACAAGAGGATGTGTCCCGACGATCT 385
QY 390 gcaaccacaacacccgttaagcttgaggcaactgcaggttttaagcctttgaggaa 449
DB 386 TCAACCCAAACACACCGTTAGACTTGGGCAACCTGCCGGTTTAAAGCTTTTGAATGAGA 445
QY 450 tgcgtgaaatatattgcaagcggtcagggatgctcgaagaacagaagtcacctaa 509
DB 446 TGCTTCTGAAAATATTGCAATCGGTAGGATATGCTGAGACAGAAATACCTTCA 505
QY 510 tctcaatacagatgcatctatcttgatggaaccaagaagttctatcttcttgggt 569
DB 506 CGTTCAACCAAGCAGCGATCTTCTATATGATGAGACCACCAAGAGGTTCTATATGGGT 565
QY 570 gacaataactatctcttgagggaagtgtggaataaagattacaagaacagtgaggtagt 629
DB 566 GACAGTTAATCTATGCAATTTGGGAATTTAGGAAAAAGTACACAAAAACAGTTGAGATAT 625
QY 630 tgaatcagaagtggttgcaaatgacataatgagtcctcaagaacacagcttaaaa 689
DB 626 AGATCTTGGAGGTGATCAGTTCATATGCGTATGAGATATCAAGAAATACTGCTAAAA 685
QY 690 tgcctcaaaagtaactgaaagagatccatatacaaaagaagctgttactccaggaa 749
DB 686 TGCCTCAAAAGTTGAGATGATGATCATATCAAGAGGTTGTACTCAAGGAAAT 745
QY 750 gaaatagacctttagttcacagttactgtgctatggaagaagaatcttcgycaga 809
DB 746 ACCATATGATCTCTATGTTCAAGATTACTTACCTTGGGTAGAGAAAGCATTCGACGACA 805
QY 810 gatttcaagtcgctgtgtctgctaactcttgcaatttagctgtggttgaggagc 869

```

```

DB 806 GATTTCGAACCTACACCTCCTGCTCTCTCTAAACCTTGTGCTTTTACGTGATTTATGGAAT 865
QY 870 atatacatctccgagcagagatagaagttctcgccccaagcttcagataacttga 929
DB 866 CTATATATATTCAGGAGAGAGGTTTAAAGCAACTGCTTACACTTGTGTCGAACCTTTAA 925
QY 930 tcaatcagaagaatgactcttaaggctcttaagaagtaagcactgtgaccttaaga 989
DB 926 TAAATGCANAAACAAATTCGTAGGCTCTTAAAGTGAACCTATCTTGTTCATATCAGAA 985
QY 990 ttgcatttgggtgagatagatgtgtgaggtggaagtgtgcaaaaaactcttct 1049
DB 986 TTGCACCTTTTGGTGAATTTGGATGTGAGAGAGAAATGACACAAACCTTTTTC 1045
QY 1050 tactatcttcttacttactctctgaagatgttggaactc---tgtaataaaccaa 1106
DB 1046 TTCTGATCTCTTCTTACCTACCTGAAATACCGGTATAGTTGATGACACACCTTAA 1105
QY 1107 tgcacaattcgccagttgatttgaagactcagcttaactagctgtgaaacaactc 1166
DB 1106 TTTCATACCTTGGCCCGGTGATATTGAGACTTAAGCTTAAGAAAGCTTGGCGTTAACTT 1165
QY 1167 tggagtgcaaaatcccaataccagatcttaagaaagacagtgatgaatagtg 1226
DB 1166 CGAGGATGCGAAATCTTACTTATTCATCTTGTATGAAGAAATGTAGCTTATGTATG 1225
QY 1227 ctgagatctgtctagtggtacacattggttgatgattgtgttgatgacattca 1286
DB 1226 CATGATCTTATATATACATATGTTGTTACTGTTGATGATTTGGTGTGATCCATTTGCA 1285
QY 1287 agaggttcaagtgagaaatgaatatatacagatgctctgtgtggaacccgcatg 1346
DB 1286 AAAGATTATCATGAGGAAGAAATGAATACCAAGATGTTATGGAAGCTGATGGCC 1345
QY 1347 tctaggaactgcacatagaagaacatatacatcttgccttaatttgagattaatgt 1406
DB 1346 TCTAGGCAATGCTGTAGAAGCCATATACCTTAAATTTGAGCATGATGTATTT 1405
QY 1407 tatttaactact 1419
DB 1406 TGTTTAACTTCT 1418

RESULT 11
E51056
LOCUS
DEFINITION
E51056 Disease resistant polypeptide, disease resistant gene, method for
impairing disease resistance to plant, disease resistance-impacted
transformed plant, and nucleotide triphosphate hydrolase.
ACCESSION
E51056
VERSION
E51056.1 GI:18629462
KEYWORDS
JP 2001017176-A/3.
SOURCE
unidentified.
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1667)
AUTHORS
Shiraiishi,T. and Furusawa,I.
TITLE
Disease resistant polypeptide, disease resistant gene, method for
impairing disease resistance to plant, disease resistance-impacted
transformed plant, and nucleotide triphosphate hydrolase
JOURNAL
Patent: JP 2001017176-A 3 23-JAN-2001;
PRESIDENT OF KYOTO UNIVERSITY
OS
Pisum sativum L. (Garden Pea)
PN
JP 2001017176-A/3
PD
23-JAN-2001
PF
02-JUL-1999 JP 1999189129
PR
TOMONORI SHIRAIISHI, IMAO FURUSAWA
PI
C12N15/09,A01H5/00,C07K14/15,C12N9/16,C12N15/00 CC
PC
key
FH
Location/Qualifiers
FT
source 1..1667
/organism="Pisum sativum L. (Garden Pea)".

```

FEATURES	Location/Qualifiers
source	1. .167
	/organism="unidentified"
	/db_xref="taxon:32644"
BASE COUNT	516 a 309 c 308 g 534 t
ORIGIN	

Query Match	56.1%;	Score 835;	DB 6;	Length 1667;
Best Local Similarity	76.0%;	Pred. No. 7.2e-186;		
Matches 1059;	Conservative	0;	Mismatches 325;	Indels 9;
				Gaps 2;

OY	30	ttgactcaagaagcatggaacttcttaattagttctcaagaccttggttcattgtaatgcc	89
Db	29	tttcccccttccacaaatggaagttccctattttaaacttattctcctttctactattttctatgcc	88
OY	90	tgcatactctcccccacatactcgcgaacaacacatcttcatactgaatcgttaagaattact	149
Db	89	TGCAATTACTTCTCCCTCCCAATACCTTAGGAAACCAACCTACTACCAAGTAAGAAAGATTTCCT	148
OY	150	cccccaaatcaaggaaccagttacatcatcagctgttataattgaagtctgttagcactgg	209
Db	149	-----AAACAACAGAGAAATTTTCCTTAGCGGCTGTAATGGATGCTGTGACACTGG	202
OY	210	aagcagaagtcacatgctacaaatttgatcgaacttagatctctccgttgaaaaaga	269
Db	203	TAGTCGATTAATGTTTACCATTTTAACGAAACTTAGACCTTTCATTAATGGCAAGC	262
OY	270	actgaagtttaagattcggttaaaccggttgaagttcaagctcgttaactccgaaga	329
Db	263	TGTCGATTAATTAATTAAGTAATACACTGGTTTGAAGTCAATGCAATTAATCCAGAAC	322
OY	330	agctgcgaagatctctgaattccactctcacaagaagcagaanaatgltctcttgagcca	389
Db	323	GGCTGCAAAATCTCTCATTCACATTTTAGAGCAAGCAAGAGATGTGTCCGACGATCT	382
OY	390	gcaaccacaacacccgcttaagcttgggcaactgcagtttaagccttttgaaggga	449
Db	383	TCAACCCCAAGACACCCCTTAGACTTTGGGCAACCTGCGGTTTAAAGCTTTGAATGAGA	442
OY	450	tgcgcgcgaataatatttgcaagcggtcgaagatattgctcagaacagaagttgccttaa	509
Db	443	TGCTTCTGAAAAAGTAATTTGGAAATCGGTAAAGGATATCTCAGAACAGAAAGTACCTTCAA	502
OY	510	ttttcattcaagatgcagtatcatctctctgtgtgaaccccaagaagttctatacttgggt	569
Db	503	CGTTCAACCAAGCAGCACTTCTTAATTAATGTGAAACCCAAAGAGTTCTTAATATGGGT	562
OY	570	gacaataactatctctcttgggaagtttggaaaaagaatttcaagaagcagttgggaatgt	629
Db	563	GACAGTTAACTATGCATTGTGGAAATTTAGGGAATAATGACAAAAAACAGTTGGAGTAAT	622
OY	630	tgaatcaggggtgggttcagttgcaaatgacatatgcagttcccaagaaacacagctaaaaa	689
Db	623	AGATCTTGGAGGTGGATCGATTCAAAAGGGGTATGCAGTATCAAAAGAAACCTCTAAAAA	682
OY	690	tgtcccaaaagtacctgaagaagagatgcatacatataaagaagctgttaattccaagaaa	749
Db	683	TGCTCCAAAAGTTGCAGATGAGATGATCATATCATCAAAAGAGTTTACTTCAAGGAAT	742
OY	750	gaatatgacctatglttccagttiaactctgcgcatgagaagaagaacattcgttcaga	809
Db	743	ACCATATGATCTTTATGTTCACAGTTACTTACACTTTGGTAAGAGAAGCATCTGCAGCCA	802
OY	810	gattttcaaggtcggctgggtgttcctgctaactccttgacatttagcttggcttgaagggc	869
Db	803	GATTTTAAAGCTCACTCTGCTCTCTCAACCTTGCTTTTAGCTGAGTAATATGAAT	862
OY	870	ataatcattccggagcagagataaggtcgtcgccccaagctcaagatcctaattga	929
Db	863	CTATATCATATTCCAGCGAAGAGATTTAAGGCACACTGCTTACACTTCTGTGCAAACTTTAA	922
OY	930	tcaattgaagaagaataagctctlaagctcttlaaagtgaatgaacactgtccctatcaaa	989

Db	923	TAAAGCAAAAAACAAATTCGTAAAGCCCTTAAAGTTGAACATATCCATTCATCCATFACACAA	982
Oy	990	ttgcacatttggttggatagatggaatcgtgtggaagtggatggatcgaataaatctttcct	1049
Db	983	TTGCACCTTTGGTGGAAATTTGGAAATGCTGGAGAGGAATGACAGAAAAACCTTTTGGC	1042
Oy	1050	tactcattcttctatctacctctctgaagatgttggatcct--tggatbaaaccaa	1106
Db	1043	TTCTTCATCTTCTTTTACCTACCTGAAGATACCGGTAATGGTTGATGCAAGCACACTAA	1102
Oy	1107	tgcgaatctgtccagttgatttgaagactgcagctcaactagcttgttaaaccaatct	1166
Db	1103	TTTTCACTTCTCGCGCGGCGATATTGAGACCTAAACCTTAAGAAAGCTTCGCGCTTAACTT	1162
Oy	1167	tgaagatgcaaaatccaatcaccagacacttataagaagaagcagttgttaatatgtg	1226
Db	1163	CGAGATCGGAATTTACTTATTCCTTTCTTGATTAAGAAAAATGACCTTCTCATATGATG	1222
Oy	1227	cttggatcttgcctcgtgtacacatctgcttgttgaattgtgtcttgatccatcca	1286
Db	1223	CATGATCTTAAATATATCATGATATGTGTACTGCTGATGGATTTGGCTTTGATCATTGCA	1282
Oy	1287	agaggttcacgttgcgaatgaatattgaataatcagaatgctcttgtggaagccgcatggcc	1346
Db	1283	AAAGATTACATCAGGGAAGGAATTAATACCAATCTATTTGGGAAGCTGATGGCC	1342
Oy	1347	tctaggaccttgcctagagaagacataatcatcttgcttaattggagatataatatt	1406
Db	1343	TCTAGGCAATGCTGTAGAGAAAGCATATACGCTTTACCTAAATTTGAGCGATGATGATTT	1402
Oy	1407	tatttaactact 1419	
Db	1403	TGTTTAAGTTCT 1415	

RESULT	12
LOCUS	E51055
DEFINITION	E51055 1697 bp DNA linear PAT 31-JAN-2002
ACCESSION	E51055
VERSION	E51055.1 GI:18629461
KEYWORDS	JP 2001017176-A/2.
SOURCE	unidentified.
ORGANISM	unidentified.
REFERENCE	1 (bases 1 to 1697)
AUTHORS	Shiraishi, T. and Furusawa, I.
TITLE	Disease resistant polypeptide, disease resistant gene, method for imparting disease resistance to plant, disease resistance-imparted transformed plant, and nucleotide triphosphate hydrolase
JOURNAL	JP 2001017176-A 2 23-JAN-2001; PATENT; JP 2001017176-A 2 23-JAN-2001; PRESIDENT OF KYOTO UNIVERSITY
COMMENT	OS Pisum sativum L. (Garden Pea) PN JP 2001017176-A/2 PD 23-JAN-2001 PF 02-JUL-1999 JP 1999189129
PR	
PI	TOMONORI SHIRAISHI, IWAO FURUSAWA
PC	C12N15/09, A01H5/00, C07K14/415, C12N9/16, C12N15/00 CC
FH	Key Location/Qualifiers
FT	1. 1697
FT	source Location/Qualifiers
FEATURES	Location/Qualifiers
Source	1. 1697
BASE COUNT	535 a 308 c 306 g 548 t
ORIGIN	

Query Match	56.0%;	Score 833.4;	DB 6;	Length 1697;
Best Local Similarity	76.0%;	Pred. No. 1.7e-185;		
Matches 1058;	Conservative 0;	Mismatches 326;	Indels 9;	Gaps 2;
QY 30	ttgagctaagccatgacttcttaatgcttcacgaccttctgtgtatgtatgac	89		
Db 12	TTAACCTTACACAAATGGAGTCTCTTTAACTATACACTTTTCTATTTCTATGCC	71		
QY 90	tgtatctctctctcccaatatctcggaaacaacatctcctatgaatcgtaatat	149		
Db 72	TGCAATCACTTCTCTCCCAATTTAGGAACAACCTTACACAGTACAGTAAGAAATTTTCT	131		
QY 150	ccccaaatatcagaaaccagatcacatcaagctgttatattgagctggtgaacg	209		
Db 132	-----AAAACAGAGGAATTTCCCTTACGCTGTAGTATTCGATCCGGTAGCACTGG	185		
QY 210	aagcagatgcatagtctacaatttgaatgaagaacttagatctctcccttgaaacg	269		
Db 186	TATGTCCATTCATGTTTACCATTTTAAACGAACTTAGACCTTCTTCATATTGGCAAGG	245		
QY 270	acttgagtttatgattcggttgaacccggtttggtcatcgctgtcattccctga	329		
Db 246	TGTGCGAGTATTATTAATAATTAACACTTGTTGATGTATACGCTTAATAATCCAGAAC	305		
QY 330	agctcagagatctctgattccacttcataaagaagcaagaanaatgtgtctcgt	389		
Db 306	GGCTGCAAAATCTCTCATTTCCACTTTTAAAGCAGACGAAAGATGTCCCGACGATCT	365		
QY 390	gaaacccaacacacccggttaagcttggggcaactgcagtttaagcctttgaagg	449		
Db 366	TCAACCCCAAGACACCCGTTTAGACTTTGGGCAACTCCGGTTTAAAGCTTTGGAATGAGA	425		
QY 450	tgtcgtctgaaaaatatactcaagcgtgcagagatagtctcagcaagaagtgcc	509		
Db 426	TGCTTCTGAAAAAGATATTGCAATCGGTAAAGGATATGCTGAGCAACAGATACCTTAA	485		
QY 510	tgttaaatagatgcagtatctatctcttgaatgaacccaagaagttcttactt	569		
Db 486	CGTTCAACGACGACGAGTTTCTATTAATGATGGAACCAAGAGGTCTTATCATATGGGT	545		
QY 570	gacaattaaactctctcttgagggaatttggaanaagettccaagaagaagtgga	629		
Db 546	GACAGTAACTATTCGATTGGGAATTTTGGGAAAAAGTACACAAAAAGTTGGAGTAT	605		
QY 630	tgaatcaggaagtggtgcagtgcaatgacatatgcgtctcaagaagcaacgctaa	689		
Db 606	AAGATCTTGGAGGTGATCATGTTCAATGGCGTATCCATCAAAAGAAACTGCTTAAAA	665		
QY 690	tgcctcaaaagacactgaagagagatccataaagaagaagcttgaatcccaagg	749		
Db 666	TGCTCAAAAGTTGCAAGATGGAGATGATCCATCTCAAGAAAGTTGATCAAGGGGAT	725		
QY 750	gaaatatgaccttattgtcacaagttaacttgcgtatgaaagagaagactcgtga	809		
Db 726	ACCATATATGATCTTTATGTTCAAGTTACTTACACTTGTGAGAGAGATOTCCAGCGA	785		
QY 810	gatttcaagtcgctggtgtctgtctaactccttgaactttagcgttgaatgggc	869		
Db 786	GATTTTGAAGCTCACTCTGCTTCTCTTCAACCTTGGCTTTTACCTGGATTTAATGGAAT	845		
QY 870	ataatacatctcggagcagagatgaagtctgcggcccgactcaagatctaactg	929		
Db 846	CTATGCATATTCAGGGAGAAAGATTTAAGCAACATGCTTACACTTCTGTGCAAACTTAA	905		
QY 930	tcaatgcagaaagatagctcttaaggctctaaagtgaatgacactgtccatcaga	989		
Db 906	TAAATGCAAAAACAAATTCCTAAGGCTCTTAACTTCAACATATCCATGTCCATTCAGAA	965		
QY 990	ttgacactttgggtggaatggaatggttgagggtggaagtggtcaaaaaaatctt	1049		
Db 966	TTGACCTTTGGTGGAAATTTGGAAATGGGTGGAGAGGAAATGGACAGAAAACCTTTTGC	1025		
QY 1050	tacttcaatcttctataactctctctcgaagaatgttggagatct--tgtataa	1106		

	Db	1026	TTCCTCATCTTTCTTACTGTAAGATACCGGTATGGTTGATGCACAACACCCTAA	1085
	OY	1107	tgccaaaattcgtccagtlgatttgaagacatgcagctaacatgacttglttaaaaaaatct	1166
	Db	1086	TTTTCAATTCTTCGCCGCCGTCCATATTGAGACTAAAGCTAAGAAGACCTTGCCGGTTAAACTT	1145
	OY	1167	tgaagatgcaaaaatccaataaccagaactctttatgagaagaagacagtgttbaatatgtgtg	1226
	Db	1146	CGAAGATGCGCAAAATCTACTTATCCATTCTTGATGAAGAAAAATGTACTTCATATGATG	1205
	OY	1227	cttgaactctgtctcacgtgtacacattgcttgtlbtgaagatttgcttccttatccattca	1286
	Db	1206	CATGAGATCTTATATATACGATGATGGTTACGTACGTGTGATGGAFTTGGCTTATCATCATTGCA	1265
	OY	1287	aaggatgtacagtggtggcgaattgaatatcatcgatgcctctgttgaagccgatggcc	1346
	Db	1266	AAGAGATTACTATCAAGGGAAGAAAATTGAATACCAAGATGCTATTGTGGAMACCTGCATGGCC	1325
	OY	1347	tctaagaacctgcacataagaacaatalcatcatcttgcccataattggaagatlaatgtatt	1406
	Db	1326	TCTAGGCAATGCGCTGAGAACCACTATCACGCTTACCTAAATTTGAGCGATTGATATTT	1385
	OY	1407	tatttaaactact 1419	
	Db	1386	TGTTTAAGTTTCT 1398	
	RESULT	13		
	EJ1054	LOCUS		
	DEFINITION	EJ1054	1368 bp DNA linear	PAT 31-JAN-2001
	DESCRIPTION	Disease resistant polypeptide, disease resistant gene, method for imparting disease resistance to plant, disease resistance-impaired transformed plant, and nucleotide triphosphate hydrolase.		
	VERSION	EJ1054.1	GI:18629460	
	KEYWORDS	JP 2001017176-A/1.		
	SOURCE	unidentified.		
	ORGANISM	unclassified.		
	REFERENCE	1 (bases 1 to 1368)		
	AUTHORS	Shiraishi,T. and Furusawa,I.		
	TITLE	Disease resistant polypeptide, disease resistant gene, method for imparting disease resistance to plant, disease resistance-impaired transformed plant, and nucleotide triphosphate hydrolase		
	JOURNAL	Patent: JP 2001017176-A 1 23-JAN -2001;		
		PRESIDENT OF KYOTO UNIVERSITY NIHON NINJIN HAMBAI KYODOKUMINAI		
		RENKORI ASARI TECHN COOP		
	COMMENT	OS Pisum sativum L. (Garden Pea)		
		PN JP 2001017176-A/1		
		PD 23-JUN-2001		
		PF 02-JUL-1999 JP 1999189129		
		PR		
		PI TOMONORI SHIRAISHI, IWAO FURUSAWA		
		PC C12N15/09,A01H5/00,C07K14/415,C12N9/16,C12N15/00 CC		
		FH key		
		FT source		
	FEATURES	Location/Qualifiers		
	source	1..1368	/organism='Pisum sativum L. (Garden Pea)'	
		/db_xref='taxon:32644'		
	BASE COUNT	422 a 260 c 275 g 411 t		
	ORIGIN			
	Query Match	55.8%; Score 830.4; DB 6; Length 1368;		
	Best Local Similarity	76.3%; Pred. No. 8,8e-185;		
	Matches 1049; Conservative	0; Mismatches 316; Indels 9; Gaps 2;		
	OY	43 atggaactttaattagctcattgacaccttggtcatgtaattgaagccgtcatctctcc	102	
	Db	1 ATGGAGTTCCTTATAACTATACCTTCTTACATTTTCTATGTCGTCGACATCACTTCC	60	

DB	Accession	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	1026	TTCTTCATCTTTCTTTACTTACTGTAAGATACCGGATATGTTGATATCCAAAGACACACCTTAA	1368 bp	linear	31-JAN-2001							
QY	1107	ttccaaaatttcgtccagttgatttgaagacgtcagctaaactagcttctgttaaaaaaact	1368 bp	linear	31-JAN-2001							
Db	1086	TTTTCATCTCTTCGGCGGCTCATATTGAGACTAAAGCTAAAGAACCTTGGCGGTTAAACTT	1368 bp	linear	31-JAN-2001							
QY	1167	tgaagatgcaaaaattccaaatracccagactctttatgagaagaacagctgttbaatatgtgtg	1368 bp	linear	31-JAN-2001							
Db	1146	CGAGGATGCGCAAAATCTACTTATCCATTCTTGATTAAGAAAAATGTACTTCATATATGATG	1368 bp	linear	31-JAN-2001							
QY	1227	cttgaactctgtctcagctgtgacacattgctctgttgaagatttgctcttatccattca	1368 bp	linear	31-JAN-2001							
Db	1206	CATGAGATCTTATATATATACGATATGTTGATCTGCTTGATGGAFTTGGCTTATCATCATGCA	1368 bp	linear	31-JAN-2001							
QY	1287	agaggttaacagctgtgacgaattgaatatcagagatgcctctgttgaagccgacatggcc	1368 bp	linear	31-JAN-2001							
Db	1266	AAAGATTACATCAAGGGAAGGAAATGTAAATCAAGATGCTATTGTGGAMACCTGCATGCGC	1368 bp	linear	31-JAN-2001							
QY	1347	tctagaagcagccatgaagaacaatcatcatcttgcccaatttgaagagattatgtat	1368 bp	linear	31-JAN-2001							
Db	1326	TCGAGGCAATGCGCTGAGAACACCATATCAGCTTTACCTAAATTTGAGCGATTGATATTT	1368 bp	linear	31-JAN-2001							
QY	1407	tatttaaacact 1419	1368 bp	linear	31-JAN-2001							
Db	1386	TGTTTAAGTTCT 1398	1368 bp	linear	31-JAN-2001							
RESULT	13											
LOCUS	E51054											
DEFINITION	E51054		1368 bp	DNA	linear	31-JAN-2001						
ACCESSION	E51054											
VERSION	E51054.1		GI:18629460									
KEYWORDS	JP 2001017176-A/1.											
SOURCE	unidentified.											
ORGANISM	unclassified.											
REFERENCE	1 (bases 1 to 1368)											
AUTHORS	Shiraishi, T. and Furusawa, T.											
TITLE	Disease resistant polypeptide, disease resistant gene, method for											
JOURNAL	impairing disease resistance to plant, disease resistance-impaired											
COMMENT	transformed plant, and nucleotide triphosphate hydrolase.											
	Patent: JP 2001017176-A 1 23-JAN-2001;											
	PRESIDENT OF KYOTO UNIVERSITY NIHON NINJIN HAMBAI KYODOKUMINAI											
	RENOKKI ASahi TECHNO GLASS CORP											
	OS Pisum sativum L. (Garden Pea)											
	PN JP 2001017176-A/1											
	PD 23-JAN-2001											
	PF 02-JUL-1999											
	PR 1999189129											
	PI TOMONORI SHIRAISHI, IWAO FURUSAWA											
	PC C12N15/09, A01H5/00, C07K14/415, C12N9/16, C12N15/00 CC											
	PF key											
	FT Location/Qualifiers											
	FT 1. .1368											
	FT Location/Qualifiers											
	FT 1. .1368											
	FT Location/Qualifiers											
	FT 1. .1368											
	FT Location/Qualifiers											
	FT 1. .1368											
	FT Location/Qualifiers											
	FT 1. .1368											
	FT Location/Qualifiers											
	FT 1. .1368											

Query Match 55.8%; Score 830.4; DB 8; Length 1557;
 Best Local Similarity 76.8%; Pred. No. 8.7e-185;
 Matches 1043; Conservative 0; Mismatches 306; Indels 9; Gaps 2;

Qy 65 tacccttggtgctatgtaagctgtctatctctccccaatatctcggaacaaca 124
 Db 1 TCACCTTTTCTACTATTTCTTATGCTGCAATCCTCCCAATACCTAGGAACAAC 60

Qy 125 tctcatgaatcgtlaaatalattactccccaaaatcaggaaccagatcatcatagcgtg 184
 Db 61 TACTCAGCAGTAGAAGATTTTCCT-----AAACACAGAGAAATTTCTCTTACGCTG 114

Qy 165 ttatattgagctgctgtagcactggaagcagaagtcacatgctacaatttgatcagaact 244
 Db 115 TCGTATTCGATGCTGGTGGACACTGGTAGTTCATTCATGTTTACCATTTTAACAGAACT 174

Qy 245 tagatctctcccttgtaaaaaaggaactgagtttatgattgctggttaaccgggttga 304
 Db 175 TAGATCTTCTTCATATTGGCAAAAGGTGCGAGTATTATTAATGAATACACCTGGTTTGA 234

Qy 305 gtccatacgtctgaatcctgaagaagctcagaatcctctgattcccacttcaagaag 364
 Db 235 GTTCATACGCTAATTAATCCAGAACAGGCTGCAAAATCTCATTTCCACTTTTACAGCAAG 294

Qy 365 cagaaatggtgttccctgctgtgagccagcaaccccaacaccccgtaagcttgaggcaactg 424
 Db 295 CAGAAAGATGTGCTCCCGACAGCTTTCACCCAGAACACCCGTTAGACTTGGGCAACTG 354

Qy 425 caggtttaaagcttctggaagggaatgctgtgtaaaatatacttcaagcggtaaggata 484
 Db 355 CCGGTTTAAGGCTTTTGAATGAGATGCTTCTGAAAAGATTTTGCAATCGGTAAAGGATA 414

Qy 485 tgcctgaacagagaagtgccttaattgtcaatcagaatgcatactctcttbatgaa 544
 Db 415 TGCAGAGCAACAGAGTACCTTCAACGTTCAACAGACGAGTTCATATTAATGATGAA 474

Qy 545 cccaagaaggtctatctcttggtgagcaataactctccctgggggaagttaggaanaa 604
 Db 475 CCCAAGAAAGGTTCTTATCTATGCTGAGACGTTAACTATGCAATTTGGGAAATTTGGGGATA 534

Qy 605 gattacaagaacagtgaggagtagtctaggaagtggtgcagtcgaatgacatatg 664
 Db 535 AGTACACAAAAACAGTTGAGTATATGATCTTGAGAGTGATCAGTTCATATGGCGATG 594

Qy 665 cagctcaagaagaacagctcaaaaatgctccaaaagtactcgaagagagagatccataca 724
 Db 595 CAGTATCAAGAAAGACGTGTAATAATGCTCAAAAAGTTGGCAGATGAGATCCATAC 654

Qy 725 taaagaagcttgtaaccaggaagaataatgaccttattgttcaagttacttgcgt 784
 Db 655 TCAAGAAGGTTGTACTCAAGGAATACCATATGATCTCTATGTTCCACAGTACTTACACT 714

Qy 785 atggaagaagaacattcgtagcagaagatttcaagtcgctggtggttctgctaatccct 844
 Db 715 TCGGTAAAGAAAGATCTCGAGAGATTTTGAAGCTCACTCTCTCTCTCAACCTT 774

Qy 845 gcaatttaagctgcttgatgaggscatalacatctccgagaagagtttaaggtctgcg 904
 Db 775 GCCTTTTAGCTGATTTAATGAATCTATACATATTCAAGAGAGGTTTAAAGCAACTG 834

Qy 905 ccccaagcttcaagatcatcattgaatcaatgacgaagaagatagctcttaagctctaaag 964
 Db 835 CTTCACCTTCTGCTGCAAACTTAAATTAATGCAAAAACAAATTCGTAAAGCTTTAAAT 894

Qy 965 tgaatgaccttgctccctacgaatctgacatttggtgggatatggaatggtgaggtg 1024
 Db 895 TGAATCATCTCTGTCATTCAGAAATTTGCACTTTGGTGAATTTGGAATGATGAGAGAG 954

Qy 1025 gaagtgtaaaaaaatcttccctactactactcttcttactactctctgaagaatgtg 1084
 Db 955 GAAATGACAGAAACCTTTTGGCTTCTTCACTTTTCTTTTACCTACCTGAAGATACCG 1014

Qy 1085 ggaactcttgt---gaataaaaccnaatgcnaaaattcgctcagttgattggaactgcag 1141

Db 1015 GTATGGTTGATGCCAGACACCTTAATTTCACTTCGGCGGTCGATATTGAGACTAAAG 1074

Qy 1142 ctaactagcttgtaaaacaacatctgagagatgcaaatcccaatcccaagctctatg 1201
 Db 1075 CTAAAGAAGCTTGCGCTTAACCTTGAGAGATGCGAAATCTACTTATTCATTTCTTGATA 1134

Qy 1202 agaaagcagtggtgaatatagtgtgcttggaatcttgctcagtgtaacacattgctgtg 1261
 Db 1135 AGAAAATTTAGCTTCATATGATGATGATGATGATTCATATATACATGATGCTTACTGCTG 1194

Qy 1262 atgattggtcttgatccaattccaagagttacagtgccgaatgaatgaatcagg 1321
 Db 1195 ATGATTTGGTCTTGATCCATTCGAAAGATTACATCAGAGGAGAAATTAATTAACAAAG 1254

Qy 1322 atgctcttgtagaagccgaatgagctcttaggcactgacataagaagaatatacatgac 1381
 Db 1255 ATGCTATTGTGGAAGCTGCATGCGCTCTAGGCAATGCTGTAGAAAGCATATCAGCTTAC 1314

Qy 1382 cttaattgagagattaatgatttatttaactact 1419
 Db 1315 CTAAATTTGAGCGATGATGATATTGTTTAAAGTTTCT 1352

RESULT 15
 AF139807
 LOCUS
 DEFINITION
 Dolichos biflorus nod factor binding lectin-nucleotide
 ACCESSION
 AF139807
 VERSION
 AF139807.1 GI:4868374
 KEYWORDS
 SOURCE
 ORGANISM
 Dolichos biflorus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Dolichos.
 1 (bases 1 to 1608)
 Etzler,M.E., Kalsi,G., Ewing,N.N., Roberts,N.J., Day,R.B. and
 Murphy,J.B.
 A nod factor binding lectin with apyrase activity from legume roots
 Proc. Natl. Acad. Sci. U.S.A. 96 (10), 5856-5861 (1999)
 2 (bases 1 to 1608)
 Etzler,M.E., Kalsi,G., Ewing,N.N., Roberts,N.J., Day,R.B. and
 Murphy,J.B.
 Direct Submission
 Submitted (31-MAR-1999) Molecular and Cellular Biology, University
 of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA
 FEATURES
 source
 1..1608
 /organism="Dolichos biflorus"
 /db_xref="taxon:3840"
 /rname="type="root"
 1..1608
 /gene="LNP"
 50..1438
 /gene="LNP"
 /function="apyrase"
 /codon_start=1
 /product="nod factor binding lectin-nucleotide
 phosphohydrolase"
 /protein_id="A031285.1"
 /db_xref="GI:4868375"
 /translation="MMWVWKTKSMSEFLLITFLRSLPKISSQVYVGSILLNRKI
 LPMDELITSAVIFPDGSSGSRVHYVNFQNDLHLHGDLETKIRKGLSSVADKP
 EKAESLILPLEAEADVPEELRPKTPILKIGATAGLILDGAEEKILQAVEMERNR
 SLSVQPDVAVIDGTQEGSYLWTVNYLLGKIGKFTYGVYIDLGASVQMAVAYS
 RNTAKNAPKPPQGEDPYMKLVLGKRYDLVHSYLRGNDARVIRFKTGGAAVSPC
 LLAGYEDIIYRSESYNIYQVGSFNPENCRLALQILRLNEPCSHENCTPGIYDNG
 KSGGKILVYTSAFYRSSFYGVFPNPKNRPLDEPTAKOKACISITPEEASSTPNV
 EKDKLPVCFDIYQYTLILVDGGLDPDEITVABIEIYQDAIVEITAMPILGTAIRATIS

BASE COUNT 471 a 317 c 346 g 474 t
 ORIGIN

SLPKFNRLMYFI*
 Query Match .55.48; Score 824.6; DB 8; Length 1608;
 Best Local Similarity 76.7%; Pred. No. 2e-183;
 Matches 1049; Conservative 0; Mismatches 309; Indels 9; Gaps 3;

Qy 51 cttaattgctcattgaccttgggttcattatgctcctctactcttccccaata 110
 Db 85 CTTCCTACTCCTCATCACTCTTCTACTCTCTCATCTGCAAAACTTCTTCTCCATA 144
 Qy 111 tctcgaaacacacattctcgaatcgtgaagataattactcccaaaatcaggaaccgt 170
 Db 145 TGTGGGAGACAGTACTACTACTAATCATCGTAAGTACTTCCCAAC---CAGAACTCCT 201
 Qy 171 tacatcatacgtctgtatattgactgctgaacacgtgaagcagagtcacatgctacaa 230
 Db 202 TACCTCTTACCTGTCTCATCTTGTGATGCTGTAGCTGTGGAGTCTGTCCATGTCTCAA 261
 Qy 221 tttagatagaacttagatctccctccgttgaaacgaacttgagtttatgtatcgt 290
 Db 262 TTTTGACGAGACTAGATCTCTCTGACATGTGGCAATGACCTCGAGTTTACAAAAAAT 321
 Qy 291 taaacccggttgatgctacacgtcgtcaatccctgaagaagctgcagatctcgtatcc 350
 Db 322 CAACCCCGTTGAGCTATACGCTGATAGCCGTGAAAAGCTGAGATCTCTATTC 381
 Qy 351 actctaaagaagaagaanaatgctgctcgtgagcagaaccccaacacacccgttaa 410
 Db 382 ACTTTTGAGAGAGAGCTGAGATGTGTCTGAGAGACTGCACCCAGACACCCCTTAA 441
 Qy 411 gcttggggcaactgcaggttgaagccttggagggaatgctgctgtaaaatatactga 470
 Db 442 GCTTGGGGCAACAGCAGGTTTGAGGCTTGTGATGGGATGCTGCTGAAAAGATATTGCA 501
 Qy 471 agcggctcaggatagctcgaacagaagtgcccttaatttcaatcagatcagatc 530
 Db 502 AGCGGTATGGGAATGTTTCAGAGAACAGATTCCTTGACGCTTCAACTGATGCGATATC 561
 Qy 531 tattcttgatggaacccaagaagttcttactcttgggtgacaattactatctctgg 590
 Db 562 TGTATTGATGGAACCCAGAAAGTTTCTTACTTATGCTTACGTTACTATCTGTAG 621
 Qy 591 gaagtgggaaaaagatttacaagaacgctgggagtagtcatcaggagtggtcagt 650
 Db 622 AAAGTTGGGAAAGAGTTTACAAAACTGTGGAGTGTATGATCTTGGAGGTCTTCAGT 681
 Qy 651 gcaaatgacatgcagctcgaagaacacagctaaaatgctccaaagtacactgaag 710
 Db 682 TCAATATGCTTATGTGTCTCAAGAAATACAGCTTAAAAATGCCCAAAACCAACCAAG 741
 Qy 711 agaggatcacatacacaagaagctgtactccagggaagaataatgaccttattatca 770
 Db 742 AGAGATCATATACATGAAGAAGCTGTACTCAAGGAAAGAAATATGACCTTATGTTCA 801
 Qy 771 cagttacttgctcagatggaagaagcaatctcgtagagagatttcaaggtcgtcgtg 830
 Db 802 CAGTTACTTGGGTTATGTGTAAACGACGACGCTTAAAGATTTTAAAGACCACTGATG 861
 Qy 831 ttctgcaatccttgacatttagctgcttgatgagggcatatatatatatcggagacaga 890
 Db 862 TGCTGCTAGTCTTGTCTTATGTCGACGCTATGAAAGATATATACAGATATTCGGGAATC 921
 Qy 891 gataaagctcgc 950
 Db 922 GTACAATATCTATGATGCCACTTGTGGTCAACTTATATGATGCTGACCTAGCTCT 981
 Qy 951 taagcctttaaagtgaatgcacactgtccctcatcagaatgcacttgggtggatg 1010
 Db 982 TCAGATTCTCAGATTGAATGAGCCATGTTCATGAAATGCACTTGTGGTGGATATG 1041

Qy 1011 gaatggtgagtggaagtggtcacaanaaattcttcccttactcatcttccattacc 1070
 Db 1042 GGATGCTGGAAGAGAAAGATGACAGAAAAACCTTGTGTACTTACGTTTCTACTATAG 1101
 Qy 1071 ctctgaagatgltgggactcttgtgaataaaccacatgccaanaattcgtccagtgtatt 1130
 Db 1102 GTCCTCTGAGGTTG---TTTGTCTACTCTCCCAATTCCAAAATVCCCGCTGTGATTT 1158
 Qy 1131 gaagactgcagctaaactagcttgttaanaaactctgaagatgcaaatccaatccc 1190
 Db 1159 TGAACCTGCAGCTAAACAAGCTTGTAGTTTAAACATTCAGAGAAAGCAATCCATTTTC 1218
 Qy 1191 agatcttatgagaagaacagtggtgaatagtggtgcttgatcgtctgaatgctacac 1250
 Db 1219 AAATGT---TGAGAAAGATTAACCTTCATTTGTATGCTGTGATTCACATACAGTATAC 1275
 Qy 1251 attgctgtgaltgagatttgctctgacatctcaagaaggttcaagtggcgaatgaat 1310
 Db 1276 ATTGCTGTGTGATGATGATTTGGCTGTAGATCCAGAGCAGAGAGATTACAGTGGCAGAAAGAT 1335
 Qy 1311 tgaataagaatgctcttgggaagccgcacatggtccttaggcactgcatagagaacat 1370
 Db 1336 TGAATATCAAGATGCGCATTTGTGAAACAGCATGGCTCTAGAGAACTGCCATAGAACCAT 1395
 Qy 1371 atcatcaltgcctaaatttgagagataatgatttatttaacta 1417
 Db 1396 ATCATCTTTGCTTAATTTAATCGTCTAATGATTTTATTCATAAGCA 1442

Search completed: June 28, 2002, 03:00:29
 Job time: 5319 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2002, 04:00:36 ; Search time 37.72 Seconds

(without alignments)
1176.916 Million cell updates/sec

Title: US-09-657-631-10

Perfect score: 2382

Sequence: 1 LHWTKAMDFELISLMTFVEML.....GTAIEAIISSLPKFERLMYFI 462

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1675.5	70.3	455	2	nucleoside triphos
2	1182.5	49.6	454	2	aprase (EC 3.6.1.
3	540	22.7	556	2	probable guanosine
4	478.5	20.1	479	2	hypothetical prote
5	442.5	18.6	518	2	guanosine-diphosph
6	431.5	18.1	483	2	hypothetical prote
7	421.5	17.7	510	2	lymphoid cell acti
8	393	16.5	516	2	probable nucleosid
9	379	15.9	405	2	hypothetical prote
10	379	15.9	572	2	probable nucleotid
11	378.5	15.9	557	2	hypothetical prote
12	375	15.7	485	2	hypothetical prote
13	343.5	14.4	630	2	hypothetical prote
14	329	13.8	508	2	hypothetical prote
15	236.5	9.9	1052	2	7A19.33 protein -
16	160.5	6.7	628	2	hypothetical prote
17	124	5.1	739	2	nucleoside-triphos
18	121	5.1	739	2	probable cytosol a
19	115.5	4.8	1417	2	beta-lactamase cla
20	115.5	4.8	1417	2	probable adhesin f
21	115	4.8	1058	2	kinasin-related pr
22	110.5	4.6	880	2	beta-glucosidase (
23	109	4.6	625	2	general secretion
24	108.5	4.6	421	2	hypothetical prote
25	107.5	4.5	891	2	DNA mismatch repai
26	106	4.5	2154	2	hypothetical prote
27	105.5	4.4	330	2	probable dihydroxy
28	104.5	4.4	349	2	gcpE protein (Impo
29	103.5	4.3	417	2	conserved hypothet

30	103.5	4.3	421	2	AF1686	hypothetical prote
31	103	4.3	565	2	T39863	zinc finger protei
32	103	4.3	1383	2	T06091	hypothetical prote
33	102.5	4.3	629	2	E64610	outer membrane pro
34	102	4.3	705	2	H83823	polysphosphate kina
35	101.5	4.3	544	2	F97170	uncharacterized co
36	101.5	4.3	890	2	A97750	DNA mismatch repai
37	101	4.2	531	2	S75607	glucose-6-phosphat
38	101	4.2	1069	2	AF1930	hypothetical prote
39	101	4.2	1519	2	S41525	major ring-forming
40	100	4.2	446	2	S26965	NADH oxidase - Ent
41	100	4.2	537	2	G83083	probable binding p
42	100	4.2	1204	2	C75015	probable pyrolysin
43	99.5	4.2	342	2	AF3301	dihydroxyacetone k
44	99.5	4.2	498	2	G97062	glycerol kinase, G
45	99.5	4.2	677	2	D97335	hypothetical prote

ALIGNMENTS

RESULT 1

S48859

nucleoside triphosphatase precursor, chromatin-associated - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999

C:Accession: S65147; S48859

R:Hu, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.

Plant Mol. Biol. 30, 135-147, 1996

A:Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chroma

A:Reference number: S65141; MUID:96197404

A:Accession: S65147

A:Molecule type: mRNA

A:Residues: 1-455 <HS2>

A:Cross-References: EMBL:Z372743; NID:9563611; PIDN:CA83655.1; PID:9563612

C:Superfamily: nucleoside triphosphatase Chromatin-associated

C:Keywords: nucleu

Query Match	70.3%	Score 1675.5;	DB 2;	Length 455;
Best Local Similarity	68.7%	Pred. No. 4.3e-115;		
Matches 314;	Conservative 64;	Mismatches 76;	Indels 3;	Gaps 2;
QY	7	MDFLISLMTFVEMLMPAIISSOYLNNILNKRILPKNOEPTVSYAVIFDAGSTSRVH	66	
DB	1	MELLILITFLFSPAITSSOYLNNILNKRIFL--KOEISSTAYVPDAGSTSRH	58	
QY	67	VYVFNQNDLPLPYENLEFEYDVKPGLSSYANPEEASLPLLEAEVNVVPSQOPI	126	
DB	59	VYVFNQNDLPLHKGVEYVYKTTPLGLSSYANPEEASLPLLEAEVNVVPSQOPI	118	
QY	127	PKVLGATAGLRLEGNAEHLIDAVDMLSNRKALVQSDAVSILDTGEGSLMTVINY	186	
DB	119	PVRLGATAGLRLEGNAEHLIDAVDMLSNRKALVQSDAVSILDTGEGSLMTVINY	178	
QY	187	LGKLGKRFKTYGVVDLGGSSVQMTYAVSRNAPKAPVPEEDDYIKKLVYLOGKKYD	246	
DB	179	ALGNLAKKRTTYGVVDLGGSSVQMTYAVSRNAPKAPVPEEDDYIKKLVYLOGKKYD	238	
QY	247	VYHSYLRYGREAIFRAEIFVAGGSANPCILAGFGAYTYSAGAEYKVASGASNLQCKR	306	
DB	239	VYHSYLHFRERASRAEILKLTSPSPNCLLAGNGIYTSGEFKATVATSGANFKCKN	298	
QY	307	IALKALKVAPCPYONCTGGIYNGGSGGOKKFLTSSFFIYSEVG--IFVKKPAKR	365	
DB	299	TIRKALKLVPCPYONCTGGIYNGGSGGOKKFLTSSFFIYSEVG--IFVKKPAKR	358	
QY	366	PVDLKAALKLACKTNEDKSKYPDLVYKDSVEYVCLDLYVYVTLVLDGSGDPPQVNY	425	
DB	359	PVDLKAALKLACKTNEDKSKYPDLVYKDSVEYVCLDLYVYVTLVLDGSGDPPQVNY	418	
QY	426	ANEIEYQDALVFAAMPGLGTAIEAIISSLPKFERLMYFI 462		


```

Db 207 SLGGDLPLKTTGIVELGGASAOVTFVSS-----PMPE-----PSRTISFGNVITNLYKSH 256
QY 250 SYLRKYR-----EAPRAEIPKVVAGGSANPCIIAGFD-GAYTVS-- 286
Db 257 SFLHGGQMAAHDKLWGLSLSRDHSNAGVPEPTREKIF-----TDPCAPRGYNLDANTOKHL 310
QY 287 -----GAEKVVS-APASGSNLRNOCRRALKALKVNAPCPYQNCFTFGGINNGGSGSQKLF 341
Db 311 SGLLAEEERLSDSPVAGGNTYSOCSRALTTILO-----DNGN-----RLII 350
QY 342 LTSSFFYYLSEDDVGFVFNKPNKIRPVDLKTAAKLACTNLLEDAKSKYDPDLYEKDSVEVYC 401
Db 351 ILAGSFLFFELG-----EKAMLSNMISAGERCGEDMSKLRVKDPSLHEEDLLRY-C 402
QY 402 LDLYVYVYTLVLDGCFDLPQFEVYANLEIEYODALVEAMPLGTATIE 447
Db 403 FSSAAYVSLLDHDTLGI-PLDERIGVYANQAGD--IPIDMALGAFIQ 445

```

RESULT 9
E86276
hypothetical protein F14L17.1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: E86276
R:Proteolysis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E86276
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <STO>
A:Cross-references: GB:AE005172; NID:g7262666; PIDN:AAF43924.1; GSPDB:GN00141
C:Genetics:
;Map position: 1

Query Match	15.9%;	Score 379;	DB 2;	Length 405;
Best Local Similarity	27.6%;	Pred. No. 4.4e-20;		
Matches 116;	Conservative 71;	Mismatches 161;	Indels 72;	Gaps 18;

[illegible]

Qy 389 PDLYEKDSVEYVCLDIYVYITLLVDGEG--LDPRQEVTVANEIEYODALVEAMPGLGTAI 446
| : | | | : : | | : : : | | | |
Db 329 PTTKDKYLHRY-CFSSAYIISMHDSLGVALDD-ERIKAKASKAGKEN--IPDDWALGAFI 3844

```

RESULT 10
T40856
probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40856
R:Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21952
A:Accession: T40856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-572 <RAM>
A:Cross-references: EMBL:AL121783; PIDN:CAB57847.1; GSPDB:GN00068; SPDB:SPCC11E10.05C
A:Experimental source: strain 972h; cosmid c11E10
C:Genetics:
A:Gene: SPDB:SPCC11E10.05C
A:Map position: 3

```

Query Match	15.98;	Score 379;	DB 2;	Length 572;
Best Local Similarity	28.28;	Pred. No. 7.4e-20;		
Matches 136;	Conservative 68;	Mismatches 176;	Indels 102;	Gaps 21

Oy	49	VTSAVVEDASTOSRVIVNEFDNLD-----LLP-VENEL-----EFDYSQPKLSS	95
Db	2	VRRKGIFFIDASSSSRLIITYSMOYDITDSSLSDKYKKLPILETGTGGDGKMSLKVQPGISS	61
Oy	96	YAANDEEAES-LIPLKEAENVVPVSOQPTPVKLCGTAGLRLEGNMAENILQAVRDM	154
Db	62	FANPCKHVKKKHLELDFEAAHAI PKDVHKETPVFLSATAGMRLLGDAONKILISHACRY	120
Oy	155	LSNSRALNVO--SDAVSILDTGQSGYLMVITNLLKLGKRFRTKVGVVDLGGSVQMT	212
Db	122	IKKNDFDIPICMSNIRKVIDKAGMGMVLTATNLTLLTLEKDJSTYGFIDMGASVQIA	189
Oy	213	YAVSBNTAKNAPKVVEGEDPIYKKLVL-----QGRKIDLYVHSYLRIG-REAFR---	266
Db	182	FELPPSOQLK-----YKDISIVHILONGCOOLEVPLFTWTWLGCAANEAYRVL	231
Oy	261	-----AEIFKYAGGSANPCIIAGDEGATVYSGAEYKVAPASGSNLMQCRITALKALKVN	315
Db	232	GLLIESENGKKGNTLSDPCSIRG--RTYDDIGIEF-----AGTDLQCLKLTNYNLNKD	286
Oy	316	APCPYONCTFGGIWNGGSGQKNLFTLTSFYLSEDV--GIFVNRPMAKIRPDLKTA	377
Db	285	KPCSMDCPNFGDISLIPVDVFANTEFVGSEFWYTTNDVDFMGSGYHPPNF--YRKVD---	339
Oy	373	AKLACKTYLBEAKSKSYPDLYEKD-----SVEYVCLDLVYVYTLVLVDG-	415
Db	340	---EYCGEWEETMLSR---LYNKKELTPSTDENKLEKLCFKASWALNVTLHEGCDVPKSNYSS	394
Oy	416	-----GL-----DPEOEVYVAEIEYODALVFAAMPGLTAIEAIS--LPKFERILMY	460
Db	395	NDAKDGLSVIRAHYSPF---TSLKIE-----RTEVSMTLGGVLLYASNQGLAKPEVANY	444
Oy	461	FI 462	
Db	448	YM 449	

RESULT 11
T16696
hypothetical protein ROT64_4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
R:Accession: T16696
R:Miller, N.

submitted to the EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid R07E4.

A:Reference number: Z18561

A:Accession: T16696

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-557 <M1>

A:Cross-references: EMBL:039652; NID:g1049390; PID:g1049394; PIDN:AAA80403.1; CESP:R07E4

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:R07E4.4

A:Introns: 39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2

Query Match

Best Local Similarity 15.9%; Score 378.5; DB 2; Length 557;
27.7%; Pred. No. 7.7e-20;
Matches 125; Conservative 70; Mismatches 190; Indels 67; Gaps 17;

```

QY 51 STAVIFDAGSTGSRVHYNFQNDLLPEVNELEFYDS-----VKPGLSSYANPEEA 104
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 43 STGVICDAGSTGTRLEFVNMISTDSELIQLEPVIYDNKPVMKISPGLSTFGTKPAQAA 102
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 105 ESILPLKEAEVNVVPSOQPNTPVKLGATAGLRLEG-----MAENILOAVRDMLSRS 159
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 EYLRLMLAEAEHRIPEKRPITPVFIATAGRLIPDEYVLIGQKAVLKNLRNKLPRKIT 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 160 ALNVOSDAVSIIDGTOEGSYLWTVINYLGLKGRFT-----KTVGVDLGG 206
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 163 SMQVLEKEHRIIEGMEGIYSWIAVNAIGKFNKATLIDPGETPAHAKQKTCGLIDMG 222
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 207 GSVQMTYAVSRN---TAAKNAKVPREG--EDPYIKVLVLOGKKYDLVHASYLRG-REAFR 260
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 ASAQIAFEIPTDSTFSSTINVENINLGCREDLSL-----FKYKLFVTFPGYGVNEGIR 275
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 261 -----AEIFVYAGSA--NPCILAGFDGAYTSGAEKVSAPASGSLNCRKIALKAL 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 276 KYEHMLSKLDQNGCTVIQDDCMPLNLKTYLTLENGENFR--KGTGMWNTCSSEVKKL 333
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 313 KVNAP---CPYQ--NCTFGGIMNGGSGGQKNLFLTSSFYLSBDV-GIFVKNKNAKIRP 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 334 NPESSEYCKAEAKCYGAAPAPASIPLSNTEWGFSEYWSYTHDVLGLGQYDAENIAK 393
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 367 VDLTKAACKTNLEDAKSYPLDYKDSVEYVCLDLYVYITLLVNGFGD---EPQE 422
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 394 KTOOYCSKRNSTIOAESKKQLYPPA-DEERLRTQCFKSAWITTVLHDGFSVDKTHNRFQS 452
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 423 V-TVANLEIYODALVEAAMPGLGTAIEAISLSP 453
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 453 VSTIAGQ-----EYQWALGAMITYHMRFPF 476
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 12

T34147 hypothetical protein C33H5.14 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000

C:Accession: T34147

R:Bradshaw H.; Stellyes, L.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C33H5.

A:Reference number: Z21482

A:Accession: T34147

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-485

A:Cross-references: EMBL:U41007; PIDN:AAA82272.1; CESP:C33H5.14

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C33H5.14

A:Introns: 22/1; 83/1; 120/2; 167/1; 269/3; 399/3

C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 15.7%; Score 375; DB 2; Length 485;

Best Local Similarity 26.6%; Pred. No. 1.1e-19;

Matches 122; Conservative 73; Mismatches 183; Indels 80; Gaps 18;

```

QY 38 RKILPKNOEVEVTSYAVIFDAGSTGSRVHYNFQNDLLPEVNELEFYDSVKR 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 RSVYETKNN---IKGYICDAGSSGTRLFVYTLKPLSGGLNIDTLHESE-PVKKVTP 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 GLSSYANPEEAASLPLKEAEVNVVPSOQPNTPVKLGATAGLRLEGMAENILOAV 151
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 70 GLSFGDKPEQVEYELPLRFAEHIPIYQELGETDLLIFATAGMRLLPQAQDAIINKL 129
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 152 RDMLSNRNALNVOSDAVSIIDGTOEGSYLWTVINYLGLKGRFTKTVGVVDLGGSVQ 211
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 QNGKSTYALRVSDSNIRITIDGANEGLYSWIAVNYLGRDKNDKNSKVMIDMGASVOI 189
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 212 TVAVSRNTAK-----NAPKVEGEDPYIKKVLVLOGKKYDLVHASYLRG---EA 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 190 AFEIANKESEYNGCNVYEINLGSLETNED-----KYKIYSTTFLGTGANGGLK 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 259 FRAEIFVYAGSANPCILAFD---GAYTYSG-AEKVSAPASGSLNCRKIALKALV 314
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 YENSLV-SGNSNDSCSPRLNRLIGFTVNGGEMDVCIAGVSLIGD-----KA 289
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 315 NACPYNCTFGGIMNGGSGGQKNLFLTSSFYLSBDV---IFVKNKNAKIRPVDLK 370
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 QPSCPNTCGLRVNIAVSVALSTVOLXGESEYVITTSNGSGGEYHYQKFTDVKRY--- 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 371 TAAKLACKTNLED-----AKSKYPLDYKDSVEYVCLDLYVYITLLVNGFGDLP---PQ 421
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 347 -----CQKDMNDIODGFKNRFPNA-DIERLTGNCFKAAWTSVLDGFNDKTRHLEQ 399
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 422 EV-TVANLEIYODALVEAAMPGLGTAIEAISLSPKFERL 458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 400 SVLKINGE-----EMQWALGAMLYHSKDL-KENLL 428
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 13

S50463 hypothetical protein YER005W - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 23-Mar-2001

C:Accession: S50463

R:Dieckrich, F.S.

submitted to the EMBL Data Library, December 1994

A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lamb

A:Reference number: S50433

A:Accession: S50463

A:Molecule type: DNA

A:Residues: 1-630 <DIE>

A:Cross-references: EMBL:U18778; NID:g603592; PID:g603597; GSPDB:GN00005; MIPS:YER005

C:Genetics:

A:Gene: MIPS:YER005W

A:Map position: 5R

Query Match

Best Local Similarity 14.4%; Score 343.5; DB 2; Length 630;
24.3%; Pred. No. 3.4e-17;
Matches 112; Conservative 78; Mismatches 177; Indels 93; Gaps 16;

```

QY 52 YAVITFDAGSTGSRVHYNFQNDLLPEVNE-----ELFFDYSVKGLSSY 96
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 10 FQIVIDAGSSGSRHIFWQDTESLHATNDOSGSIILQSVPHIQEKDWFRLNPGSSSF 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 97 AANPEEAESLI-PLTKAEVNVVPSOQPNTPVKLGATAGLRLEGMAENILOAVRDM 155
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 70 EKKPDATKSHKPLPLDRAKNITPESHWSCVFLQATAGMKLLPQDIQSSILDLQGL 129
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 156 SNRSALNQ--SDAVSIIDGTOEGSYLWTVINYLGLG-----KLGRFTVGVVDLGG 206
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 KHPAEFLVEDCSAQIQVIDGETBGLYGMNLGYLGHNDVNPVSDHF--TFPFMDMG 187
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 207 GSVQMTYAVSRNTAKNAKVPBEGDP-----YIKKVLVLOGKKYDLVHASYLRG---REA 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 188 ASTQIIFA-----PHDSGEIARHDDIATIFLRSVNGDLQKMDVFEVSTWLGAGANARR 242
OY 259 FRAEIFKVGAGS-----PCIIAGFDGATYSGAEKVASAPASGSLNOC 304
Db 243 YLAQILNTPENTINDYENDDFSTRNLNDCMPRGSSSTDEFFDTJFIH--AGSGVEYEC 299
OY 305 RKIALKALKVNAPCPYONCTPGIWMGSGGSKNLFTSSPFYIYSEDVGIYFNPKNAKI 364
Db 300 TKSITPBLKKNPCDDEPCLFENGVAHPRIDFANDKFIGTSEWYIYAND----VFKLGGEY 355
OY 365 RPVDLTKTAACKLACKTN-----LEDAKSKY--PDLYEKDSVEVYCHDLVYVYTLVDGFG 416
Db 356 NPDFKFSKSLREFCNSMWTQILANSOKGYNSIPENFLKDA-----CFKGMVNLIIHEGRD 411
OY 417 LDP-----FQEVYVANEIYEDALVEAMPYG 443
Db 412 MPRIDVDAENVDRPLFQSEKVEE-----RELSTWLG 444

```

RESULT 14

```

C86276
7A19.33 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: C86276
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakono, H.
A:Author: Rooney, T.; Rowley, D.; Sakono, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C86276
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-508 <STO>
A:Cross-References: GB:A8005172; NID:95080800; PIDN:AAD39310.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

```

Query Match 13.8%; Score 329; DB 2; Length 508;

Best Local Similarity 25.7%; Pred. No. 2.9e-16;

Matches 122; Conservative 88; Mismatches 179; Indels 86; Gaps 21;

```

OY 4 TKAMDF-LTSLMTFVFMPLPAITSSOYLGNIL--MNRKILLPKNOEPTVSAYVFDAG 59
Db 29 TYSITFIYACTIALGL-----FTGYSLRSGRRKRVSL-----HTSVIIDG 73
OY 60 STGSRVHYVNFQNDLLLEVENLEFEYDSVK--PGLSSYAANDEEAESLIPLLKAENV 117
Db 74 SSGTRVHVGGRLESQPVFEGCEENYASLKLSGLSAYADNEGSESTELVEYAKKR 133
OY 118 VEPSQOPNTPVKGATAGRLLEGNAENILQAVRDMLSKRSALNQSDAVSLDGTQEG 177
Db 134 VHKGLKSKDIRIMATAGKRLLELPVOEQILDVTRRVLRS--SGFDRDMASVISEILN 192
OY 178 SYLWVNIINLLGKGRFTKTYGVNDLGGGYOMTYAASRNNAKNAKPYEGEDPYIKKL 237
Db 193 FQDLMAV-YMLGIL-----LIMRSVRLLEVTV-----STELVP--SEFSRTL 231
OY 238 VLQGRYDLYVHSYLRGGEAFRAEIPKVAAGSA-----NPCILAGF----- 279
Db 232 AVGNVSYNLYSHSLDFGDDAQEKLESILYNSAANSTGEGVDPDCIPKGYILLETNLOK 291
OY 280 --DGATYSGAEKVASAPASGSLNOCRKIALKAL-KVNAPCPYONCTPGIYNN--GGG 333
Db 292 DLPGFLADKG-KFTATLQAG--NFSECRSAAFMLOEKGKCTYKRCISIGSIFTPLNLOS 349

```

```

OY 334 GSGOKNLFTSSPFYIYSEDVGIYFNPKNAKIRPVDLTKTAACKLACKTNLEDAKSKYPDLYE 393
Db 350 FLATENEFHTSKFFGGEKEML-----SEMIILAKRRCGSENSKIKVYKTYTFPD 398
OY 394 KDSVEVYCHDLVYVYTLVDGFG--LDPROEYVANEIYEDALVEAMPPLGTAI 446
Db 399 ENLIRY-CRSSAYIISMLHDSLGVALDD--ERIKYAKAGEED--IPLDMALGAFI 449

```

RESULT 15

T04439

hypothetical protein T18B16.150 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: T04439

R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes,

submitted to the Protein Sequence Database, April 1998

A:Reference number: T15359

A:Accession: T04439

A:Molecule type: DNA

A:Residues: 1-1052 <BEV>

A:Cross-References: EMBL:AL021687

A:Experimental source: cultivar Columbia; BAC clone T18B16

C:Genetics:

A:Map position: 4

A:Intons: 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3

A>Note: T18B16.150

Query Match 9.9%; Score 236.5; DB 2; Length 1052;

Best Local Similarity 22.6%; Pred. No. 5.1e-09;

Matches 108; Conservative 69; Mismatches 164; Indels 137; Gaps 20;

```

OY 54 VIFDAGSTGRVHY-----NFDONLLDLEVENL-----EYD--SVKPGISRA 97
Db 525 LVYIVSITGTTRAVVYOASINYYKDDSLPYMKSITFGISNRKGRAYDRMETEPGDKLV 584
OY 98 ANPEEAESLIPLKEAENVVPSQOPNTPVKLGATAGRLLEGNAENILQAVRDMLSN 157
Db 585 NNRTGLKTAIKPLQWAEKOIPKNAHRTISLEYATAGVRRLRPADSSNLTGVMGSLA- 643
OY 158 RSALNQSDAVSLDGTQGSYLMWTINY--LLGRLGRFTKTYGVNDLGGGSVOMTYA 214
Db 644 KSPFCRREWVKLIISCTEAEYFGWTALNYQTSMLGALPKK--ATFGALDGGSSLGVTFE 701
OY 215 VSRNTAKNAKPYEGEDPYIKLVLOGKKYDLYVHSYLRG-REAF----- 259
Db 702 NEERTHNETN-----LNLRIGSVNHHLISAVSLAGLNDARDRSVYHLKRLPNV 751
OY 260 -RAEIFKVGAGSANPCIIAGFDGAYTSGAEYVSAAPASG-----NLNOCRKI 307
Db 752 NKSDDLIEGLEKMKHPCLNSGYNQYICSOCASSVQGGKKKSSVSLKLVGAPMNGEC--- 808
OY 308 ALKALKVNAPC-----PYONCTPGIWMGSGGSKNLFTSSPFYIYSEDVGIYFNK 359
Db 809 --SALKNAKNAKPYEGYPRH-----GQ-----FYANS--GEFV-- 838
OY 360 PNAKIRPVDLTKTAACKLACKTNLEDAKSKYPDLYE-----DSVEYCHDLVYV 407
Db 839 -----VYREFNLSAESISLDDVLEKGRFEDKAMOVAKRVSPOPIEQCFRAPHYI 889
OY 408 YTLVDGFGDPRQOEYVANEIYEDALVEAMPPLGTAI---EATSS--LPKPERL 458
Db 890 VSLIRGGLYITDKQIIGSGSI-----TWTLGVALLESKALSLTGLKSYETL 938

```

Search completed: June 28, 2002, 04:06:52

Job time: 376 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2002, 02:12:41 ; Search time 19.5 Seconds
(without alignments)
578.699 Million cell updates/sec

Title: US-09-657-631-10

Perfect score: 2382

Sequence: 1 LHWTKAMDPLISLMTFVFM.....GTAIEALISLPKFERLMYFI 462

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*

2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/Backfill1est.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1675.5	70.3	455	US-09-240-639-10	Sequence 10, Appl
2	1182.5	49.6	454	US-09-240-639-11	Sequence 11, Appl
3	491.5	20.6	428	US-09-608-285A-3	Sequence 3, Appl
4	491.5	20.6	428	US-09-608-285A-5	Sequence 5, Appl
5	491.5	20.6	428	US-09-240-639-6	Sequence 6, Appl
6	491.5	20.6	428	US-09-240-639-9	Sequence 9, Appl
7	486.5	20.4	428	US-09-608-285A-7	Sequence 7, Appl
8	472	19.8	405	US-09-608-285A-25	Sequence 25, Appl
9	463	19.4	454	US-09-240-639-2	Sequence 2, Appl
10	463	19.4	484	US-09-608-285A-27	Sequence 27, Appl
11	460	19.3	465	US-09-240-639-8	Sequence 8, Appl
12	443	18.6	473	US-09-240-639-12	Sequence 12, Appl
13	421.5	17.7	510	US-08-930-921-1	Sequence 1, Appl
14	392.5	16.5	471	US-09-608-285A-60	Sequence 60, Appl
15	315.5	13.2	529	US-09-240-639-4	Sequence 4, Appl
16	251.5	10.6	148	US-09-240-639-17	Sequence 17, Appl
17	224.5	9.4	154	US-09-240-639-14	Sequence 14, Appl
18	213	8.9	153	US-09-240-639-13	Sequence 13, Appl
19	197.5	8.3	153	US-09-240-639-15	Sequence 15, Appl
20	196	8.2	150	US-09-240-639-16	Sequence 16, Appl
21	106.5	4.5	1027	US-08-446-137B-2	Sequence 2, Appl
22	97	4.1	872	US-08-766-014-2	Sequence 2, Appl
23	95.5	4.0	905	US-09-074-658-70	Sequence 70, Appl
24	95	4.0	2004	US-08-375-709-15	Sequence 15, Appl
25	95	4.0	2004	US-08-752-929-15	Sequence 15, Appl
26	95	4.0	2004	US-09-090-793-9	Sequence 9, Appl
27	94	3.9	587	US-08-931-608A-4	Sequence 4, Appl

28	93	3.9	487	US-08-961-083-42	Sequence 42, Appl
29	92.5	3.9	761	US-09-625-188-14	Sequence 14, Appl
30	92	3.9	450	US-08-861-464-2	Sequence 2, Appl
31	92	3.9	450	US-08-396-001-2	Sequence 2, Appl
32	92	3.9	450	US-09-323-433A-2	Sequence 2, Appl
33	92	3.9	934	US-08-457-176-2	Sequence 2, Appl
34	92	3.9	934	US-08-457-175-2	Sequence 2, Appl
35	92	3.9	934	US-08-709-784-1	Sequence 2, Appl
36	92	3.9	934	US-09-651-656-3	Sequence 3, Appl
37	92	3.9	1112	US-08-714-402-2	Sequence 2, Appl
38	92	3.9	1161	US-09-327-536-2	Sequence 2, Appl
39	91	3.8	1403	US-08-387-942C-3	Sequence 3, Appl
40	88.5	3.7	2237	US-08-354-973-1	Sequence 1, Appl
41	88	3.7	564	US-08-425-843-8	Sequence 8, Appl
42	88	3.7	565	US-08-425-843-3	Sequence 3, Appl
43	87.5	3.7	1457	US-08-652-971-3	Sequence 3, Appl
44	87.5	3.7	1457	US-08-449-644-1	Sequence 1, Appl
45	87.5	3.7	1457	US-08-087-244A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-240-639-10
; Sequence 10, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischaut, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; FILE REFERENCE: 9598-066
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 455
; TYPE: PRT
; ORGANISM: P. sativum
US-09-240-639-10

Query Match	70.3%	Score 1675.5;	DB 4;	Length 455;
Best Local Similarity	68.7%	Pred. No. 2.2e-165;		
Matches 314;	Conservative 64;	Mismatches 76;	Indels 3;	Gaps 2;
QY	7	MDPLISLMTFVFMIMPAISSOYLGNINIMNRKILPKNOEPTSYAVIFDAGSTGRVH	66	
DB	1	MELLILITFLFSPMPATISSOYLGNINILTSKIFL--KOEISSTAVYFDAGSTGRH	58	
QY	67	VYNFDONIDLPLENELEFYDSVYKGLSSYANPEBAESLIPLKEANVYVPSQOPNT	126	
DB	59	VYHFNQMLDLHIGKVEYEVYKRTPLSSYANPEQAKSLIPLEQADVDVDDIQPT	118	
QY	127	PVKLGATLRLLEGNAENILQAVDMLSNSALWVDAYSILDQEGSLMTWTIVY	186	
DB	119	PVRLGATLRLRLNGASERILQSVDMLSNSTFVNQPDVAYSIDQEGSLMTWTIVY	178	
QY	187	LGKLGKRTKTYGVVDLGGGSVOMTYAVSRNTAKNAPVEGEDDYIKKLVLOGKKYD	246	
DB	179	ALGNLCKKTYKTYGVVDLGGGSVOMTYAVSKTAKNAPVADDDPYIKKLVLOGKIPYD	238	
QY	247	YVHSYLRGREAIFRAIFVAGSANPCILAGFDGAYVYSGAEYKVSAPASGSLNQCK	306	
DB	239	YVHSYLRGREAIFRAIFVAGSANPCILAGFDGAYVYSGAEYKVSAPASGSLNQCK	298	
QY	307	IALKALKVAPCPYONCTGGIWMNGGGGOKNLFTSSFYLYSEVG--IPNKKPAKIR	365	
DB	299	TIRKALKVAPCPYONCTGGIWMNGGGGOKNLFTSSFYLYSEVG--IPNKKPAKIR	358	
QY	366	PVDLKAATLACKTNLEDAKSKYPDLYEKDSVEYVCLDVLVYVYTLVDGFLDPFOEYTV	425	

```

Db 359 PVDIEFKAKEACALNFEDAKSTYFPLDKKNVASYCMQDLYQYVLVDGFGDLPLQKITS 418
OY 426 ANEIEYOALVEAAMPICGTATIAISLSLKEFELAMFI 462
Db 419 GKEIEYODALVEAAMPICGNAVEAISALPKFERLMYFV 455

```

RESULT 2
US-09-240-639-11

```

; Sequence 11, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischauf, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 454
; TYPE: PRN
; ORGANISM: Solanum tuberosum
US-09-240-639-11

```

Query Match 49.6%; Score 1182.5; DB 4; Length 454;
Best Local Similarity 51.1%; Pred. No. 3.8e-114;
Matches 226; Conservative 85; Mismatches 124; Indels 7; Gaps 4;

```

OY 11 ISLMTFVFMLEPAISSOYGCNNTIMNRKILPKNOEPVSYAVIFDAGSGSVHYHNF 70
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 9 IFILALFVPLSLSKNVMAQIPLRHLHSESH---YAVIFDAGSGSVHYHNF 64
OY 71 DONIDLLPVENELEFYDSVYKRGISVYANPEEAESLIPLIKEANVPVSOQFPNPKL 130
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 65 DEKIGLPIGNINIEYFVATERGLSSYAEPPKAANSLEPLDGEVGVPELQSEPLEL 124
OY 131 GATGGLLEENAEINLOAVRMLSNRSALNOSDAVSIIDTQOESSTYLMWTINYLK 190
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 125 GATGGLLEENAEINLOAVRMLSNRSALNOSDAVSIIDTQOESSTYLMWTINYLK 190
OY 191 LGRKFTYGVVLDGGGSVOMTYAVSRNTAKNAPKVEGEPYIKKLVLOGKKYDLYVHS 250
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 185 LGRKFTYGVVLDGGGSVOMTYAVSRNTAKNAPKVEGEPYIKKLVLOGKKYDLYVHS 250
OY 251 YLRGGRAPFAEIRFKVAGGSANPCITLAGFGAYTYSGAETKVSAPASGSLNOCRIALK 310
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 244 YLNGQLAGRAEITFKASRNSNPCALEGCGYYSYGVDYKAPKPKSSMKRCRLTRH 303
OY 311 ALKNAPCPYONCTFEGGIMWGGSGGOKNLFITSSFYLLSEDDGIFVNR-PNAKIRPVLD 369
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 304 ALKNAPCPYONCTFEGGIMWGGSGGOKNLFITSSFYLLSEDDGIFVNR-PNAKIRPVLD 369
OY 370 KTAAKLACKTNLEDAKSKYDLYEKDSVEVYCDLYVYVTLVLDGFGDLDFQEVYANET 429
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 364 LNAKAVACQNTNVAIDIKSIFPKTQDR-NIPYLCMDLIEYTLVLDGGLNHNKRIYIHYD 422
OY 430 EYODALVEAAMPICGTATIAISLS 451
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 423 QYKNYLVGAAMPICGCAIDIVSS 444

```

RESULT 3
US-09-608-285A-3
; Sequence 3, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio

```

; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-608-285A-3

```

Query Match 20.6%; Score 491.5; DB 4; Length 428;
Best Local Similarity 31.3%; Pred. No. 2.3e-42;
Matches 134; Conservative 71; Mismatches 166; Indels 57; Gaps 14;

```

OY 41 LIPKNOEPVSYAVITDASTGSRVHYVFDONL-DLPEVNELEFYDSVYKRGISVYAN 99
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 38 MCPINVSASTLVGIMPDASTGTRIHVYFVOKMPQOLPI-LEGEVFDVSKPKGLSAFVQ 96
OY 100 PEBAESLIPLIKEANVPVSOQFPNPKLGTATGLRLKEGNAENITLOAVRMLSNRS 159
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 97 PKGATFVGLLEVAKDSIPRSHMKKTVPVLAITAGLRLPEHKAKALLEVEKEIF-RKS 155
OY 160 ALNVDASVSIIDTQOESSTYLMWTINYLGLKGRFTKTVGVVLDGGGSVOMTYAVSRNT 219
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 156 PELVPRKGSVSIIMDSDEGILAVTVFVLGQLHGRHOETVGTLDGASTOITFL-PQF 213
OY 220 ANKAPRVPGEDEPYIKKLVLOGKKYDLYVHSYLRKGRERF-----AEIKV 266
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 214 EYODALVEAAMPICGTATIAISLS 451
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 267 AGGSANPCITLAGFGAYTYSGAETKVSAPASGS-MLNOCRIALKLVNAPCPYONCTF 325
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 271 A-----CLPRMLAEWIMIGVYKQYGGNQBEGVEPEPCYAEVLAVNGKLLHQPBE---- 320
OY 326 GGIWNGGGSGGOKNLFITSSFY-LSEDDGIFVNR-PNAKIRPVLDKTAAKLACKTNLEDA 384
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 321 -----VORGSEFYAFSYDYDRAVDITMDIYERKGLKVEDFERKAREVC-DNLENF 369
OY 385 KSKYDLYEKDSVEVYCDLYVYVTLVLDGFGDLDFQEVYANETLEYODALVEAAMPICGT 444
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 370 TSGSP-----FLCMDLITLALDKDGF--FAUSTVL-QLTKRVNNIETGVALGA 416
OY 445 ATEAISL 452
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 417 TFLHLSL 424

```

RESULT 4


```

; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-240-639-9

```

```

Query Match
Best Local Similarity 31.3%; Score 491.5; DB 4; Length 428;
Matches 134; Conservative 71; Mismatches 166; Indels 57; Gaps 14;

```

```

QY 41 LPRKQEPVTSYAVIFDAGSTGRVHYVNFQNL-DLPEVNELEFYDVKPGLSSYAN 99
DB 38 MCPINVASSTLYGIMFDAGSTGTRIHVYTFVQKMPGOLPI-LEGEVFDVSKPGLSAFVDQ 96
QY 100 PEEAASLIPLKEAENVVPSQOPNTPVKLGATAGRLLEGNAENITLQAVRDMLSNRS 159
DB 97 PKQGAETVQGLLEVAKDSIPRSHMKRTPVYLKATAGRLLEPHKAKALLEVEKEIF-RKS 155
QY 160 ALNVOSDAVIDDQEGSYLWNTINYLKLGKRFKFTVGVVDLGGSSVQMTAVSRNT 219
DB 156 PELVKGKSVSMDGDEGLAWVTWVNLGOLHGHROETVGLDGLGASTQITFL-PQF 213
QY 220 AKNAPKEBEDPYIKKLVLQGGKDYLVHYSYLRGREAFR-----AEIKV 266
DB 214 EKTLEQTPRG---YLTFEMFNSTYKLYTHSYLGFGLKAALATLGALETGEGHTFRS 270
QY 267 AGGSANPCILAGFDGAYTYSAGAEKVSAPASGS-NLNOCRIALKALKVNAPCPYONCTF 325
DB 271 A-----CLPRMLEMWFEGVKYQYGGNOGEVGFEPCEYAEVLRVVGKLLHOPE--- 320
QY 326 GGIMNGGGSGCKNLFTSFEY-LSEVDYGFVKNPKAKIRPVDLKTAACKTNLEDA 384
DB 321 -----VORGSEYAFSYIDRAVDPTMDIDYEGKILKEDPERKAREVC-DNLENF 369
QY 385 KSKYDLYEKRSVEYCLDLYVYVTLVDGGLDPDEVTYANIEYODALVEAMPGLT 444
DB 370 TSGSP-----FLCMLDSYITALLKDFG---FADSTVL-QLTKVNNIETGMALGA 416
QY 445 AIEAISL 452
DB 417 TFFHLOSL 424

```

```

RESULT 7
US-09-608-285A-7
; Sequence 7, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Mulero, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09

```

```

; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-7

```

```

Query Match
Best Local Similarity 31.1%; Score 486.5; DB 4; Length 428;
Matches 133; Conservative 71; Mismatches 167; Indels 57; Gaps 14;

```

```

QY 41 LPRKQEPVTSYAVIFDAGSTGRVHYVNFQNL-DLPEVNELEFYDVKPGLSSYAN 99
DB 38 MCPINVASSTLYGIMFDAGSTGTRIHVYTFVQKMPGOLPI-LEGEVFDVSKPGLSAFVDQ 96
QY 100 PEEAASLIPLKEAENVVPSQOPNTPVKLGATAGRLLEGNAENITLQAVRDMLSNRS 159
DB 97 PKQGAETVQGLLEVAKDSIPRSHMKRTPVYLKATAGRLLEPHKAKALLEVEKEIF-RKS 155
QY 160 ALNVOSDAVIDDQEGSYLWNTINYLKLGKRFKFTVGVVDLGGSSVQMTAVSRNT 219
DB 156 PELVKGKSVSMDGDEGLAWVTWVNLGOLHGHROETVGLDGLGASTQITFL-PQF 213
QY 220 AKNAPKEBEDPYIKKLVLQGGKDYLVHYSYLRGREAFR-----AEIKV 266
DB 214 EKTLEQTPRG---YLTFEMFNSTYKLYTHSYLGFGLKAALATLGALETGEGHTFRS 270
QY 267 AGGSANPCILAGFDGAYTYSAGAEKVSAPASGS-NLNOCRIALKALKVNAPCPYONCTF 325
DB 271 A-----CLPRMLEMWFEGVKYQYGGNOGEVGFEPCEYAEVLRVVGKLLHOPE--- 320
QY 326 GGIMNGGGSGCKNLFTSFEY-LSEVDYGFVKNPKAKIRPVDLKTAACKTNLEDA 384
DB 321 -----VORGSEYAFSYIDRAVDPTMDIDYEGKILKEDPERKAREVC-DNLENF 369
QY 385 KSKYDLYEKRSVEYCLDLYVYVTLVDGGLDPDEVTYANIEYODALVEAMPGLT 444
DB 370 TSGSP-----FLCMLDSYITALLKDFG---FADSTVL-QLTKVNNIETGMALGA 416
QY 445 AIEAISL 452
DB 417 TFFHLOSL 424

```

```

RESULT 8
US-09-608-285A-25
; Sequence 25, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Mulero, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11

```



```
;; PRIOR APPLICATION NUMBER: 09/370,265
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: PCT/US99/16180
;; PRIOR FILING DATE: 1999-07-16
;; PRIOR APPLICATION NUMBER: 09/350,836
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: 09/273,447
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 09/244,444
;; PRIOR FILING DATE: 1999-02-04
;; PRIOR APPLICATION NUMBER: 09/122,449
;; PRIOR FILING DATE: 1998-07-24
;; PRIOR APPLICATION NUMBER: 09/118,205
;; PRIOR FILING DATE: 1998-07-16
;; NUMBER OF SEQ ID NOS: 60
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 25
;; LENGTH: 405
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-608-285A-25
```

```
Query Match      19.8%; Score 472; DB 4; Length 405;
Best Local Similarity 31.9%; Pred. No. 2.2e-40;
Matches 128; Conservative 66; Mismatches 151; Indels 56; Gaps 13;
```

```
QY 41 LLPKQEPYTSYAVIFDAGSTGSRVHYVNFQNL-DLHVENLEFYDSVKGSLSSYAN 99
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 38 MCPINVASSTLYGIMFDAGSTGTRHVTYVQKMPQGLPI-LEGEYFDSVKGLSAFVQ 96
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 100 PEEAESILPLKEANENVPVSOQNTPYKLGATAGLRLLEGNAENITIOAVRDMLSNRS 159
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 97 PROGAFTVGLLEVAKDSIFPRSHMKKTPYVLTAKTAGLRLLEPKAKALLFEYKEIR-RKS 155
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 160 AINVODASJLDGTOEGSYLWVTINYLKLGKRFKTVGVVVDLGGSGVOMTYAVSRMT 219
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 156 PELVPRGSYSIMDSDEGLTAWVTNFTLQHGHRQETVGLDGGASTQTFL--PQF 213
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 220 AKNAKRVPEGEPIYKLVLOKCKDLYHSLRYGRERF-----AEIFV 266
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 214 EHTLQTPRG---YLTFSEFMNSTYKLYTHSYLGLKAKARLATLGALETGDTGHTFRS 270
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 267 AGGSANPCLLAGEFDGAYVYTSAGAEYKVSAPASGS-NLNQCKTALKKLVNACOPYQNCIF 325
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 271 A-----CLPRWLEAMWIFGVKTYQGGNQEVEGPEPCYAEVLRVRSKLLHQPEE---- 320
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 326 GGIMNGGSGGQKNLFTSSFFY-LSEDVGIFVKNPKNKRIRPVDLKTAKTNLEDA 384
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 321 -----VGRSFFYAFSYYYDRAVDYDMIDYEKGGLKVEDPERKARVC-DLLENF 369
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 385 KSKYEDLVKESVEYCDLYVYVYLLVDGFGLDPPQEVTV 425
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 370 TSGSP-----FLCMDLSYITALKKGDFG--FADSTV 398
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 9
US-09-240-639-2
;; Sequence 2, Application US/09240639
;; Patent No. 6350447
;; GENERAL INFORMATION:
;; APPLICANT: Chadwick, Brian Paul
;; APPLICANT: Frischaut, Anna-Maria
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
;; FILE REFERENCE: 9598-066
;; CURRENT APPLICATION NUMBER: US/09/240,639
;; PRIOR FILING DATE: 1998-01-29
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 2
;; LENGTH: 456
;; TYPE: PRT
```

```
;; ORGANISM: Homo sapiens
US-09-240-639-2
```

```
Query Match      19.4%; Score 463; DB 4; Length 456;
Best Local Similarity 31.3%; Pred. No. 2.4e-39;
Matches 131; Conservative 70; Mismatches 156; Indels 62; Gaps 13;
```

```
QY 52 YAVIPAGSTGSRVHYVNFQNDLLPVENLEFYDSVKGSLSSYANPEEAESILPIL 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 74 YGIMFDAGSTGTRHVTYVQKMPQGLPI-LEGEYFDSVKGLSAFVQ 132
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 112 KEAENVPVSOQNTPYKLGATAGLRLLEGNAENITIOAVRDMLSNRSALNVOSDAVSIL 171
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 133 DVAKQDIPDPFKATPVLKATAGLRLLEGKQKLOKVEKF-KASPLVDDDEVSTIM 191
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 172 DGTQEGSYLWVTINYLKLGKRFKTVGVVVDLGGSGVOMTYAVSRMTAKNAPV----- 226
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 192 NGTDEGVSAWITINFLTGSLSKTPGSSVGMLDGGSGTQIAF-----LPRVGTIQ 242
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 227 --PEGEDPYIKLVLOGKRYDLYVHSYLRKGRFAELFKVAGS-----ANPCL 276
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 243 ASPPG---YLTLRPMFNRTYKLYSYSYLGILMSARLAILGVGEQPARDKELVSPCLIS 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 277 AGFDGAYVYTSAGAEYKVSAPASGSNLNQ-CRKIALKAL--KVNAPCPYQNTFCGIMNGG 333
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 PSKSGMEHAETVYRYSQKAAASLHELCAARYSEVLQNRVHTEVKKHYD----- 351
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 334 GSGQKNLFTSSFFYLSDEVGIFVKNPKNKRIRPVDLKTAKTNLDAKSKYDLYE 393
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 352 -----YAFSYYDDLAAGVGLDAEKGSLLVGDPEIAKAYCPT-LETPQSSP----- 399
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 394 KDSVEYCDLYVYVYLLVDGFGLDPPQEVYANLEFYODALVEAMPRTGATAEALSSL 452
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 400 -----FSCMDLYTV-SLLIQERGFPRSKVLKILTRKIDN-----VEJSMALGALFHYIDSL 448
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 10
US-09-608-285A-27
;; Sequence 27, Application US/09608285A
;; Patent No. 6335013
;; GENERAL INFORMATION:
;; APPLICANT: Ford, John
;; APPLICANT: Mulero, Julio
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
;; FILE REFERENCE: 28110/36570
;; CURRENT APPLICATION NUMBER: US/09/608,285A
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 09/583,231
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 09/557,800
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: 09/481,238
;; PRIOR FILING DATE: 2000-01-11
;; PRIOR APPLICATION NUMBER: 09/370,265
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: PCT/US99/16180
;; PRIOR FILING DATE: 1999-07-16
;; PRIOR APPLICATION NUMBER: 09/350,836
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: 09/273,447
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 09/244,444
;; PRIOR FILING DATE: 1999-02-04
;; PRIOR APPLICATION NUMBER: 09/122,449
;; PRIOR FILING DATE: 1998-07-24
;; PRIOR APPLICATION NUMBER: 09/118,205
;; PRIOR FILING DATE: 1998-07-16
;; NUMBER OF SEQ ID NOS: 60
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 27
```

LENGTH: 484
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-608-285A-27

Query Match 19.4%; Score 463; DB 4; Length 484;
 Best Local Similarity 31.3%; Pred. No. 2.6e-39;
 Matches 131; Conservative 70; Mismatches 156; Indels 62; Gaps 13;

QY 52 YAVIFDAGSGSRVHYVNPQNDLLEFVENELEFYDSYKGLSSYAANPEBAESLIPLL 111
 DB 102 YGIMFDAGSGSRVHYVNPQNDLLEFVENELEFYDSYKGLSSYAANPEBAESLIPLL 160
 QY 112 KEENAVYVQSQPPTPKVATAGLRLLLEGNAENIIQAVRDMLSNSALNVQDAVSL 171
 DB 161 DVAKODIPFPFKATPVLVATAGLRLLLEGNAENIIQAVRDMLSNSALNVQDAVSL 219
 QY 172 DGTQEGSYLWVTINYLGLGKRFKTVGVVDLGGGSVQMTYAVSRNTAKNAPKV----- 226
 DB 220 NGTDEGSAMITINFLGSLKTPGSSVGMIDLGGSGTQIAF-----LPRVEGTIQ 270
 QY 227 --PEGEDPYIKKLVLOGKKDYVHSTLYRGREAFRAIFKVVAGS-----ANPCIL 276
 DB 271 ASPPG--VLTALRMFRRTYKLYSYSLGLMSARLAILGVEGQPAKDKKELVSPCLS 327
 QY 277 AGPDGATYSGAEKYSAPASGSLNQ--CKRIKAL--KVNAFCPYQNCPTFGSINNNGG 333
 DB 328 PSFGEWEHEAVTIRVSGOKRAASLHELCAARVSEVLQNRHTEEVKHDF----- 379
 QY 334 GSGQKNLFLTSFYLSSEVDYGFVKNKPNKIRPYDLTKAAKLNLEDAKSKYPDLYE 393
 DB 380 -----YASYYDLAAGVGLDAEKGSVLVGDDEFIAKKYGRV--LEHQPOSSP----- 427
 QY 394 KDSVEYVCLDLYVYTLVYVGFGLDPQEVTVANEITYQDALVEANPLGTALTAISL 452
 DB 428 -----FSCMDLTV--SLLOEFGFPRSKVLKTRKIDN---VETSWALGAIHFYIDSL 476

RESULT 11

US-09-240-639-8
 ; Sequence 8, Application US/09240639
 ; Patent No. 6350447
 ; GENERAL INFORMATION:
 ; APPLICANT: Chadwick, Brian Paul
 ; APPLICANT: Frischauf, Anna-Maria
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
 ; FILE REFERENCE: 9598-066
 ; CURRENT APPLICATION NUMBER: US/09/240,639
 ; CURRENT FILING DATE: 1998-01-29
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 465
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-240-639-8

Query Match 19.3%; Score 460; DB 4; Length 465;
 Best Local Similarity 30.1%; Pred. No. 5e-39;
 Matches 132; Conservative 68; Mismatches 174; Indels 64; Gaps 14;

QY 17 VFMLMPAIISSOYLGNINLRK-----ILPKNOEPTVSIAVIFDAGSGSRVH 66
 DB 8 VFMLIIAC-----VGSIVFYREQQTEWEGVFLSMCPINVSAGTFYGIIMFDAGSGTARIH 62
 QY 67 VYVNFQ-----NLDLLEPVENELEFYDSYKGLSSYAANPEBAESLIPLLKEAVNVVVSQ 122
 DB 63 VYVFOVKTACQGLPFL-----EGEIFDSVKGPIASAVDQPKQGAETVQVLELAVAKSIPRSH 118
 QY 123 QPNTPVKLGATAGLRLLLEGNAENIIQAVRDMLSNSALNVQDAVSEILDOTQSGSYLWV 182

DB 119 WERTPVVLTATAGLRLLLEGNAENIIQAVRDMLSNSALNVQDAVSEILDOTQSGSYLWV 177
 QY 183 TINYLLGKLRKRFKTVGVVDLGGGSVQMTYAVSRNTAKNAPKVPBEEDPYIKKLVLOGK 242
 DB 178 TVNFLTQGLRQGEVTVGLDGLGASTQITFL--POEKTLEGTQPRK---YLSFSEFENS 232
 QY 243 KYDLYVHSTLYRGREAFRAIFKV--AGS-----ANPCILAGPDGATYSGAEKYSAP 295
 DB 233 TFLIYHSTLYRGREAFRAIFKV--AGS-----ANPCILAGPDGATYSGAEKYSAP 292
 QY 296 ASGS--NINQCRKIALKAL--KVNAFCPYQNCPTFGSINNNGGSGQKNLFLTSFYLSSE 352
 DB 293 QEGEMGEFPCYAEVLRVYQGLKLPPEEVRGSAF-----YAFSYYYDRAD 337
 QY 353 VGIFVKNPNKIRPYDLTKAAKLNLEDAKSKYPDLYEKDSVEYVCLDLYVYVTLV 412
 DB 338 THLIDYKGGVLAIVEDERRAREVC--DNLGSFSSGS-----FLCMLITIALLK 387
 QY 413 DGFGLDPQEVTVANEIE 430
 DB 388 DGLGFARHPLTAHKESE 405

RESULT 12

US-09-240-639-12
 ; Sequence 12, Application US/09240639
 ; Patent No. 6350447
 ; GENERAL INFORMATION:
 ; APPLICANT: Chadwick, Brian Paul
 ; APPLICANT: Frischauf, Anna-Maria
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
 ; FILE REFERENCE: 9598-066
 ; CURRENT APPLICATION NUMBER: US/09/240,639
 ; CURRENT FILING DATE: 1998-01-29
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 12
 ; LENGTH: 473
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 ; US-09-240-639-12

Query Match 18.6%; Score 443; DB 4; Length 473;
 Best Local Similarity 31.7%; Pred. No. 3e-37;
 Matches 142; Conservative 62; Mismatches 164; Indels 80; Gaps 16;

QY 52 YAVIFDAGSGSRVHYVNPQNDLLEFVENELEFYDSYKGLSSYAANPEBAESLIPLL 105
 DB 48 YVIMIDAGSGSRVHYVNPQNDLLEFVENELEFYDSYKGLSSYAANPEBAESLIPLL 99
 QY 106 SLIPLLKEAVNVVVSQOPPTPKVATAGLRLLLEGNAENIIQAVRDMLSNSALNVQ 164
 DB 100 SLIPLLKEAVNVVVSQOPPTPKVATAGLRLLLEGNAENIIQAVRDMLSNSALNVQ 159
 QY 165 SDAYSILDTQEGSYLWVTINYLGLGKRFK--TYGVVDLGGGSVQMTYAVSRNTAKN 222
 DB 160 GDGVSIMGDEGEVFAWITINYLGNIGANGPKLPTAIVDFLGGGSTOI--VEEPFPI 216
 QY 223 APKVPBEEDPYIKKLVLOGKKDYVHSTLYRGREAFRAIFKVVAGS----- 271
 DB 217 NEKAVDEHKEF--DLKGDENYTLIYQSHLGTGLKEGRKNVSVLYVNAIKDKTLKCN 274
 QY 272 -----NPCI-----LAGPDGATYSGAEKYSAPASGSLNQCRRIATKALKVNA 316
 DB 275 KTHCLSSPCLPKVNATNEKVTLESKEYT--IDFIDPERSGA---QCRFLDEIILKDA 330
 QY 317 PCPYQNCPTFGSINNNGGSGQKNLFLTSFYLSSEVDYGFVKNKPNKIRPV----- 367
 DB 331 QCSPPCSFNGV-----HQPISLVRT---FKESNDIYIFSYFYDRTTRBLGMDLSPTLN 380

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2002, 02:11:25 ; Search time 49.6 Seconds
(without alignments)
1034.599 Million cell updates/sec

Title: US-09-657-631-10

Perfect score: 2382
Sequence: 1 LHWTKAMDPLSLMTFVFM.....GTAIEAISSLPKEERLWYFI 462

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_032802:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2363	99.2	496	20	AAW85686
2	1684	70.7	462	20	AAW85684
3	1676.5	70.4	455	22	AAW81952
4	1665.5	69.9	486	20	AAW85685
5	1412.5	59.3	467	20	AAW85687
6	684.5	28.7	243	21	AAW82051
7	618.5	26.0	217	21	AAW82052
8	529.5	22.2	461	22	ABB66213
9	529.5	22.2	464	22	ABB59611
10	491.5	20.6	428	21	AAV44849
11	491.5	20.6	428	22	AAW72238

12	491.5	20.6	428	22	AAW72243	Human CD39 like pr
13	486.5	20.4	428	21	AAV44850	Human CD39-L4 vari
14	486.5	20.4	428	22	AAW72240	Human CD39 like pr
15	472	19.8	405	21	AAV44851	Human CD39-L66 pro
16	472	19.8	405	22	AAW72239	Human CD39 like pr
17	463	19.4	484	22	AAW72241	Human CD39 like pr
18	456	19.1	456	22	AAW93929	Human polypeptide,
19	426	17.9	464	21	AAV70899	Protein encoded by
20	426	17.9	464	21	AAV70922	Human soluble CD39
21	424.5	17.8	474	21	AAV70900	Protein encoded by
22	424.5	17.8	474	21	AAV70923	Human soluble CD39
23	424.5	17.8	476	21	AAV70889	Protein encoded by
24	424.5	17.8	476	21	AAV70912	Human CD39-L4-2/3
25	423	17.8	473	21	AAV70901	Protein encoded by
26	423	17.8	473	21	AAV70924	Human soluble CD39
27	423	17.8	487	21	AAV70898	Protein encoded by
28	423	17.8	487	21	AAV70921	Human soluble CD39
29	421.5	17.7	439	22	AAW71918	Soluble human CD39
30	421.5	17.7	454	21	AAV70890	Protein encoded by
31	421.5	17.7	454	21	AAV70913	Human soluble CD39
32	421.5	17.7	463	21	AAV70902	Protein encoded by
33	421.5	17.7	463	21	AAV70925	Human soluble CD39
34	421.5	17.7	476	21	AAV70888	Protein encoded by
35	421.5	17.7	476	21	AAV70911	Human CD39-L4-1 pr
36	421.5	17.7	478	21	AAV70891	Protein encoded by
37	421.5	17.7	478	21	AAV70914	Human soluble CD39
38	421.5	17.7	510	17	AAW04264	Human CD39 protein
39	421.5	17.7	510	17	AAW04334	Human lymphoid cel
40	421.5	17.7	510	21	AAV70887	Human soluble CD39
41	421.5	17.7	510	21	AAV70910	Human soluble CD39
42	421.5	17.7	510	22	AAV71917	Human CD39. Homo
43	416	17.5	462	22	AAU30882	Novel human secret
44	392.5	16.5	471	22	AAW72242	Mature human CD39
45	361.5	15.2	495	20	AAV33296	Human membrane spa

ALIGNMENTS

RESULT 1	AAW85686	standard; Protein; 496 AA.
ID	AAW85686	
XX	AAW85686	
AC	19-JUL-1999	(first entry)
XX		
DE	NBP46 root lectin.	
XX		
KW	NBP46; lectin; Rhizobium; leguminous plant; transgenic plant;	
KW	nitrogen; nitrogen fixation; fertilizer.	
XX		
OS	Lotus japonica.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 234	/note="this residue is encoded by nucleotides missing from the specification, and so cannot be identified"
FT	Misc-difference 252	/note="this residue is encoded by nucleotides missing from the specification, and so cannot be identified"
FT	Misc-difference 270	/note="this residue is encoded by nucleotides missing from the specification, and so cannot be identified"
FT		
XX	W09907223-AL.	
XX	18-FEB-1999.	
XX	05-AUG-1998;	98WO-US16261.

0.

XX	DT	19-JUL-1999	(first entry)
XX	DE	BBP46 root lectin.	
XX	XX	BBP46: lectin; Rhizobium; leguminous plant; transgenic plant;	
KW	XX	nitrogen; nitrogen fixation; fertilizer.	
XX	XX	Dolichos biflorus.	
OS	XX		
XX	Key	Location/Qualifiers	
FT	Region	52..73	
FT	FT	/label= "Conserved region	
FT	FT	/note= "Conserved among various plant and animal	
FT	FT	apyrases"	
FT	Region	128..144	
FT	FT	/label= "Conserved region	
FT	FT	/note= "Conserved among various plant and animal	
FT	FT	apyrases"	
FT	Modified-site	139	
FT	FT	/note= "Potential N-glycosylation site"	
FT	Region	153..175	
FT	FT	/label= "Conserved region	
FT	FT	/note= "Conserved among various plant and animal	
FT	FT	apyrases"	
FT	Region	160..180	
FT	FT	/label= "Conserved region	
FT	FT	/note= "Conserved among various plant and animal	
FT	FT	apyrases"	
FT	Modified-site	276	
FT	FT	/note= "Potential N-glycosylation site"	
XX	XX	WO907223-A1.	
XX	XX	18-FEB-1999.	
XX	XX	05-AUG-1998; 98WO-US16261.	
XX	XX	06-AUG-1997; 97US-0907226.	
XX	PA	(REGC) UNIV CALIFORNIA.	
XX	PI	Etzler ME, Murphy JB;	
XX	DR	WPI; 1999-167136/14.	
XX	DR	N-PSDB; AAX08522.	
XX	PT	New polynucleotides encoding Nod factor binding lectins - useful	
XX	PT	for production of transgenic plants which are able to fix nitrogen	
XX	PS	Claim 7; Page 43; 57pp; English.	
XX	XX		
CC	CC	The BBP46 root lectin is instrumental in recognising and binding	
CC	CC	to nitrogen fixing rhizobial bacteria via a lectin-carbohydrate	
CC	CC	interaction. The production of transgenic plants comprising an	
CC	CC	expression cassette expressing the BBP46 root lectin is advantageous	
CC	CC	since it would mean that non-leguminous plants could fix nitrogen	
CC	CC	from the atmosphere, lessening the need for the addition of nitrogen	
CC	CC	containing fertilizer to soil. This would lead to higher crop yields	
CC	CC	where soil has been overplanted and replantment of the depleted	
CC	CC	soil with usable nitrogen. Alternatively, expression of BBP46 can be	
CC	CC	used to modulate oligosaccharide signalling in the plant. The nucleic	
CC	CC	acid sequences can be used to inhibit expression of an endogenous	
CC	CC	gene and also to suppress endogenous BBP46 gene expression.	
XX	XX		
XX	Sequence	462 AA:	

Query Match 70.7%; Score 1684; DB 20; Length 462;
 Best Local Similarity 68.5%; Pred. No. 4e-147;
 Matches 319; Conservative 68; Mismatches 71; Indels 8; Gaps 5;
 1 LHW---TKRADFLSLMTFVFMMLPAISSOYLNNNTILMKRIILPNNOEFTVSAAVIF 56

Db 1 mawvwpktsmsfll-11tllfslpklsssgyvgnsl1nhkklip-ngelltsyavif 58
OY 57 DAGSTSRHVHYVFNEDONLULPVENLEFEYDSKPGLSSTYAANPEEAESLIPLKEAEN 116
Db 59 dagssqsrhvhtfdqnl1l1hgnldlftkklpglssyadkpekkaesl1p1leaad 118
OY 117 VVPSQOPTPVKLTAGTARLEEGNAENI1QAVRDMLSNRSALNWSQSAV1LDGTOE 176
Db 119 vvpelhpktplk1gataq1r1ldgdaek1l1qavremftrns1svqpaav1dqtge 178
OY 177 GSYLWMTYNYLGLGKLRFTKTVGVVDLGGSVQMTYAVSRNTAKNAPKVEGEDPYIKK 236
Db 179 gsy1w1t1vny1l1gk1gk1t1k1t1c1t1v1d1l1g1s1v1q1m1a1v1s1r1n1t1a1k1n1a1p1k1v1e1g1e1d1p1y1l1o1g1k1k1t1l1y1l1v1d1g1t1g1 238
OY 237 LVLOGKKYLYVHSYLRGREAFAEFVYAGGSANPC1IAGFDGAVTYSAGAEYKVSAPA 296
Db 239 1v1k1g1k1y1d1l1v1h1s1y1l1r1g1n1d1a1r1v1k1f1k1t1d1g1a1s1p1c1l1a1g1y1e1d1l1y1r1y1s1g1e1s1y1n1y1p1t1 298
OY 297 SGNLNOGCKRIALKALKVNAPOYONCTFGIWNNGSGGOKMLFTSSSFYLSDEVGIF 356
Db 299 sgantfecdl1a1q1l1r1l1e1p1e1s1h1e1n1c1t1g1l1w1d1g1k1s1g1q1k1n1l1v1t1s1a1t1y1r1s1e1v1g1-1 357
OY 357 VKNPNAKIRPVDLKTAALACTNLEDAKSKYPDLYEKDSVEYVCIDL1VYVTL1VDGFG 416
Db 358 v1r1p1n1a1k1r1p1v1d1l1k1t1a1l1a1c1t1n1l1e1d1a1k1s1k1y1p1d1l1y1e1k1d1s1v1e1y1c1i1d1l1v1y1t1l1v1d1g1f1g1 416
OY 417 LDDPEVYVANEIEYODALVEAAPLGTAL1EALISSLPKFERLMYFI 462
Db 417 l1d1p1e1g1i1t1v1a1e1g1e1y1g1d1a1l1v1e1t1a1p1l1g1t1a1l1e1a1l1s1s1p1k1f1n1r1l1m1y1f1 462

RESULT 3

ID AAB81952 standard; protein: 455 AA.
XX AAB81952;
AC AAB81952;
XX
DT 25-JUN-2001 (first entry)
XX
DE Pea blight resistance protein.
XX
KW Pea; blight resistance; nucleotide triphosphate decomposition.
XX
OS Pisum sativum.
XX
PN JP2001017176-A.
XX
PD 23-JAN-2001.
XX
PF 02-JUL-1999; 99JP-0189129.
XX
PR 02-JUL-1999; 99JP-0189129.
XX
PA (KYOU) UNIV KYOTO.
XX
DR MPI: 2001-320697/34.
DR N-PSDB: AAF85679, AAF85680, AAF85681.
XX
PT New blight-resistant polypeptide useful for giving blight resistance to a plant -
XX
PS Claim 1: Page 12; 20pp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a pea protein with nucleotide triphosphate decomposing activity. The gene can be used for conferring blight resistance on a plant. The present sequence is the protein of the invention.
XX
SQ Sequence 455 AA;

Query Match 70.4%; Score 1676.5; DB 22; Length 455;

Best Local Similarity 68.7%; Pred. No. 1,9e-146;
Matches 314; Conservative 64; Mismatches 76; Indels 3; Gaps 2;

OY 7 MDPLISLMTFVFLMLRAISSQYLGN11LNKR1ILPKNOEPTSYAVIFDAGSTGSRVH 66
Db 1 m1e1l1i1k1l1l1f1l1s1m1p1l1e1s1s1y1l1g1n1l1t1s1r1f1l1--k1e1e1s1y1a1v1e1d1a1g1s1t1s1r1h1 58
OY 67 VVFNEDONLULPVENLEFEYDSKPGLSSTYAANPEEAESLIPLKEAENVVPVSOOPT 126
Db 59 v1v1f1n1e1d1o1n1l1u1l1p1e1n1l1e1f1e1y1d1s1v1k1p1g1l1s1s1y1a1n1p1e1g1a1e1s1l1p1l1k1e1a1e1n1v1p1v1s1o1o1p1t1 118
OY 127 PVKLTAGTARLEEGNAENI1QAVRDMLSNRSALNWSQSAV1LDGTOGSYLWMTYNY 186
Db 119 p1v1r1l1t1a1g1a1r1l1l1n1g1d1a1s1e1k1l1g1s1v1d1m1s1n1r1c1f1n1v1p1d1a1v1e1l1d1g1t1g1e1g1s1y1l1w1t1v1y1 178
OY 187 LGLGKLRFTKTVGVVDLGGSVQMTYAVSRNTAKNAPKVEGEDPYIKKLV1LOGKKYLY 246
Db 179 l1g1l1k1k1r1f1t1k1t1v1n1y1l1l1g1g1s1v1q1m1a1v1s1k1t1a1k1n1a1p1k1v1e1g1e1d1p1y1l1o1g1k1k1t1l1y1l1v1d1g1t1g1 238
OY 247 VVHSYLRGREAFAEFVYAGGSANPC1IAGFDGAVTYSAGAEYKVSAPASGNLNOGCK 306
Db 239 1v1h1s1y1l1r1g1r1e1a1r1a1e1f1f1v1a1g1s1a1n1p1c1l1a1g1f1d1g1a1v1t1y1s1a1g1e1y1k1v1s1a1p1a1s1g1n1l1n1o1g1c1k1 298
OY 307 IALKALKVNAPOYONCTFGIWNNGSGGOKMLFTSSSFYLSDEVG-1FVKNPNAKIR 365
Db 299 t1r1k1a1l1k1l1n1y1p1c1p1y1n1c1t1f1g1i1w1n1g1g1n1g1q1k1l1f1a1s1s1f1f1y1p1e1d1t1g1m1d1a1s1t1p1n1f1l1r1 358
OY 366 PVDLKTAAK1LACTNLEDAKSKYPDLYEKDSVEYVCIDL1VYVTL1VDGFGDLPDEYTV 425
Db 359 p1v1d1l1k1t1a1k1l1a1c1t1n1l1e1d1a1k1s1k1y1p1d1l1y1e1k1d1s1v1e1y1c1i1d1l1v1y1t1l1v1d1g1f1g1l1d1p1l1q1t1s1 418
OY 426 ANEIEYODALVEAAPLGTAL1EALISSLPKFERLMYFI 462
Db 419 g1k1e1y1g1d1a1l1v1e1a1p1l1g1t1a1l1e1a1l1s1l1p1k1f1e1r1l1m1y1f1 455

RESULT 4

ID AAM85685 standard; protein: 486 AA.
XX AAM85685;
AC AAM85685;
XX
DT 19-JUL-1999 (first entry)
XX
DE NBP46 root lectin.
XX
KW NBP46; lectin; Rhizobium; leguminous plant; transgenic plant;
XX
KM nitrogen; nitrogen fixation; fertilizer.
XX
OS Medicago sativa.
XX
FH Key Location/Qualifiers
FT MISC-difference 460 /note= "Encoded by TAA stop codon"
FT MISC-difference 470 /note= "Encoded by TAA stop codon"
FT MISC-difference 479 /note= "Encoded by TAA stop codon"
XX
PN MO9907223-A1.
XX
PD 18-FEB-1999.
XX
PF 05-AUG-1998; 98WO-US16261.
XX
PR 06-AUG-1997; 97US-0907226.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Etzler ME, Murphy JB;
XX
DR MPI: 1999-167136/14.
DR N-PSDB: AAX08528.

42

Example 3: Page 55; 57pp; English.

Example 3; Page 55; 57pp; English.

to nitrogen fixing rhizobial bacteria via a lectin-carbohydrate interaction. The production of transgenic plants comprising an

from the atmosphere, lessening the need for the addition of nitrogen since it would mean that non-leguminous plants could fix nitrogen themselves.

soil with usable nitrogen. Alternatively, expression of NBP46 can be used to modulate oligosaccharide signalling in the plant. The nucleic

gene and also to suppress endogenous NBPF4 gene expression. This DBX sequence also isolated from *D. biflorus* is also involved in

Sequence 467 AA;

Play match	59.38;	Score 1412.5;	DB 20;	Length 467;
Local Similarity	59.48;	Pred. No. 5.7e-122;		
Conservativity	68.	Mismatches 100;		
T-delta	13.	G----		

```
6 AMDFELISLMTFVEMLM-----PAISSQYLGNLIMRKILLPKNQEPV-----TSYAVIF 56
:||||| |:|:| | :|| | :| |||||
```

57 DAGSTGRHVY NFDQNDLLPVENELEFYDVKPGLSSIAANPEPAESLIPLIKEAEN 116

117 VVPVSPQNPITPVKLGATAGLRILEGNAENITLQAVRDMISRSALNVOSDAVSLIDGTGF 176

122 vipgelinrtipvkvgataglrqlegdasnrllqavsdmllkrstlkvegdavslsngne 181

182 gayqwtvrlnglgkhykstkavvdlggsgvmayalseedakapvpdpvesyte 241

```
242 mflrgkkylylvhsylrryglaaaraevlkvrsdsenpcilsgfdgytytygyqykattapp 301
```

[illegible]

Db	302	sgsftskcgnvialhlnvatcsykdcctfigivngggaggenmfvasffteradeag-f	360
Oy	357	V--NKPNAKTRPVDLTKPAKTLACTNLEDAASKYPDLYEKDSEYVCLDLVYVYFTLLVDG	414
Db	361	vdpdnaiaivrpvdfedaeavacstetkdlksvfprvkqgd-vpylcldlivgytllvdg	419
Oy	415	FGLPPEQVTPANEIERQDALVEAPMLGTAIEAIISSLPKFERLMYRI	462
Db	420	fgldpqgeitlvrtqdydsiveawpdisaleaissaipkfkltmytl	467
RESULT 6			
AC	ANG22051	ANG22051 standard; Protein: 243 AA.	
AC	ANG22051:		
DT	17-OCT-2000	(first entry)	
XX	Arabidopsis thaliana	protein fragment SEQ ID NO: 24833.	
XX	Protein identification: signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter		
XX	termination sequence.		
XX	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	23-APR-1999; 99US-0130891.		
PR	28-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	30-APR-1999; 99US-0132407.		
PR	04-MAY-1999; 99US-0132484.		
PR	05-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		
PR	06-MAY-1999; 99US-0132487.		
PR	07-MAY-1999; 99US-0132863.		
PR	11-MAY-1999; 99US-0134256.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134221.		
PR	14-MAY-1999; 99US-0134370.		
PR	18-MAY-1999; 99US-0134768.		
PR	19-MAY-1999; 99US-0134941.		
PR	20-MAY-1999; 99US-0135124.		
PR	21-MAY-1999; 99US-0135353.		
PR	24-MAY-1999; 99US-0135629.		
PR	25-MAY-1999; 99US-0136021.		
PR	27-MAY-1999; 99US-0136392.		
PR	28-MAY-1999; 99US-0136782.		
PR	01-JUN-1999; 99US-0137222.		
PR	03-JUN-1999; 99US-0137528.		
PR	04-JUN-1999; 99US-0137502.		
PR	07-JUN-1999; 99US-0137724.		
PR	08-JUN-1999; 99US-0138094.		

[illegible]

PR	09-AUG-1999;	990S-0147935.
PR	10-AUG-1999;	990S-0148171.
PR	11-AUG-1999;	990S-0148319.
PR	12-AUG-1999;	990S-0148341.
PR	13-AUG-1999;	990S-0148565.
PR	13-AUG-1999;	990S-0148684.
PR	16-AUG-1999;	990S-0149368.
PR	17-AUG-1999;	990S-0149175.
PR	18-AUG-1999;	990S-0149426.
PR	20-AUG-1999;	990S-0149722.
PR	20-AUG-1999;	990S-0149723.
PR	23-AUG-1999;	990S-0149929.
PR	23-AUG-1999;	990S-0149902.
PR	25-AUG-1999;	990S-0149930.
PR	26-AUG-1999;	990S-0150566.
PR	27-AUG-1999;	990S-0150884.
PR	27-AUG-1999;	990S-0151065.
PR	27-AUG-1999;	990S-0151066.
PR	27-AUG-1999;	990S-0151080.
PR	30-AUG-1999;	990S-0151303.
PR	31-AUG-1999;	990S-0151438.
PR	01-SEP-1999;	990S-0151930.
PR	07-SEP-1999;	990S-0152363.
PR	10-SEP-1999;	990S-0152637.
PR	13-SEP-1999;	990S-0153070.
PR	15-SEP-1999;	990S-0153758.
PR	16-SEP-1999;	990S-0154018.
PR	20-SEP-1999;	990S-0154039.
PR	22-SEP-1999;	990S-0154779.
PR	23-SEP-1999;	990S-0155139.
PR	24-SEP-1999;	990S-0155486.
PR	28-SEP-1999;	990S-0155659.
PR	29-SEP-1999;	990S-0156458.
PR	04-OCT-1999;	990S-0156596.
PR	05-OCT-1999;	990S-0157117.
PR	06-OCT-1999;	990S-0157753.
PR	07-OCT-1999;	990S-0157865.
PR	08-OCT-1999;	990S-0158029.
PR	12-OCT-1999;	990S-0158232.
PR	13-OCT-1999;	990S-0158369.
PR	13-OCT-1999;	990S-0159293.
PR	13-OCT-1999;	990S-0159294.
PR	14-OCT-1999;	990S-0159295.
PR	14-OCT-1999;	990S-0159329.
PR	14-OCT-1999;	990S-0159330.
PR	14-OCT-1999;	990S-0159331.
PR	18-OCT-1999;	990S-0159637.
PR	18-OCT-1999;	990S-0159638.
PR	21-OCT-1999;	990S-0159584.
PR	21-OCT-1999;	990S-0160741.
PR	21-OCT-1999;	990S-0160767.
PR	21-OCT-1999;	990S-0160768.
PR	21-OCT-1999;	990S-0160770.
PR	21-OCT-1999;	990S-0160814.
PR	21-OCT-1999;	990S-0160815.
PR	22-OCT-1999;	990S-0160980.
PR	22-OCT-1999;	990S-0160981.
PR	22-OCT-1999;	990S-0160989.
PR	25-OCT-1999;	990S-0161404.
PR	25-OCT-1999;	990S-0161405.
PR	26-OCT-1999;	990S-0161406.
PR	26-OCT-1999;	990S-0161359.
PR	26-OCT-1999;	990S-0161360.
PR	28-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161920.
PR	28-OCT-1999;	990S-0161992.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.

Qy	Db	Accession	Protein	Fragment	SRQ ID	NO
Qy	1	mayaipeadaatpkrpvegedsyvremylkgrkylflynshylhylaaraeilkvseds	1		60	
Qy	271	ANCIAGFPGATTTYSGAETKYSAPASGSNLMOCRKAIAKALV-NAPCPYOMCTGGSIW	329			
Db	61	mpciatlyagcykkygskafkaaspsgasldecrvainaalkvnnsictlmkctfgyvw	120			
Qy	330	NGGSGSGOKLFTTSFFYLTSEDVGFV--NKPNAKIRPVDLKTPAKLACKTNLEDAKSK	387			
Db	121	nsgyggggqgqkmtvasffidraaeg-fvdgpnvgaevrpldtekaankacmmmeegsk	179			
Qy	388	YPDLYEKDSVEYCIQDLYVVTLLVDGFLDPQFEVYANETEXODALVEAAMP LGTAIE	447			
Db	180	fprv-seedhlpyctldlyqycllvdgfglkpsqctltlvkvwkygdyaaveaawplgsaie	238			
Qy	448	AISS 451				
Db	239	avss 242				
RESULT 7						
AAG22052						
ID	AAG22052 standard; Protein; 217 AA.					
AC	AAG22052;					
DT	17-OCT-2000 (first entry)					
DE	Arabidopsis thaliana protein fragment SRQ ID NO: 24834.					
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.					
XX	Arabidopsis thaliana.					
XX	EP1033405-A2.					
XX	06-SEP-2000.					
XX	25-FEB-2000; 2000EP-0301439.					
PR	25-FEB-1999;	99US-0121825.				
PR	05-MAR-1999;	99US-0123180.				
PR	09-MAR-1999;	99US-0123548.				
PR	23-MAR-1999;	99US-0125788.				
PR	25-MAR-1999;	99US-0126264.				
PR	29-MAR-1999;	99US-0126785.				
PR	01-APR-1999;	99US-0127462.				
PR	06-APR-1999;	99US-0128234.				
PR	08-APR-1999;	99US-0128714.				
PR	16-APR-1999;	99US-0129845.				
PR	19-APR-1999;	99US-0130077.				
PR	21-APR-1999;	99US-0130449.				
PR	23-APR-1999;	99US-0130510.				
PR	23-APR-1999;	99US-0130891.				
PR	28-APR-1999;	99US-0131449.				
PR	30-APR-1999;	99US-0132048.				
PR	30-APR-1999;	99US-0132407.				
PR	04-MAY-1999;	99US-0132484.				
PR	05-MAY-1999;	99US-0132485.				
PR	06-MAY-1999;	99US-0132486.				
PR	06-MAY-1999;	99US-0132487.				
PR	07-MAY-1999;	99US-0132863.				
PR	11-MAY-1999;	99US-0134256.				
PR	14-MAY-1999;	99US-0134218.				
PR	14-MAY-1999;	99US-0134219.				
PR	14-MAY-1999;	99US-0134221.				
PR	14-MAY-1999;	99US-0134370.				
PR	18-MAY-1999;	99US-0134768.				
PR	19-MAY-1999;	99US-0134941.				
PR	20-MAY-1999;	99US-0135124.				
PR	21-MAY-1999;	99US-0135353.				

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.


```

Db      156 pflvpgsvsindgsdeglawtvtvfltgqlhghrgetvqclidlgastqtlfl--pgf 213
      | :|||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      220 AANAPVPEGEDPYIKLVLOGKKYDLVHVSYLRYGREAFR-----AEIRKV 266
      | :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      214 ektleqtprg---yltsfemfnslyklyhsylgfglkaarlalgaletegtdhtfrs 270
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      267 AGGSANPCILAGFDGAVTYTSGAEKVSAPASGS-NLNOCRKIALAKLVNAPCPYONCTF 325
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      271 a-----clprvlaeawifgvkyvgngqgevegfepeyaevlrvvgklhqpce--- 320
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      326 GGIMWNGGGSGQKNLFLTSSFYI-LSEDVGIFVNNKPNKIRPVDLTKTAACKTNLEDA 384
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      321 -----vgrgsfyafsyddravtdmdldyegkqllkvedferkarevc-dnlenf 369
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      385 KSKYPDLYEKDSVEYVCLDIYVYVTLVDGSLDFQEVYVANELEYDAILVEAAMPIGT 444
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      370 tsgsp-----flcmdlsyitalikdfig---fadstvl-qltkkvmnietgwaiqa 416
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      445 ALEAISL 452
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      417 tfhlqsl 424

```

```

RESULT  11
AAB72238
ID      AAB72238 standard; Protein: 428 AA.
XX
AC      AAB72238;
XX
DT      14-MAY-2001 (first entry)
XX
DE      Human CD39 like protein CD39-L4 amino acid sequence.
XX
KW      Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW      myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW      cerebral artery thrombosis; platelet aggregation; inflammation;
KW      apoptosis; autoimmune disorder; neurological disorder;
KW      Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS      Homo sapiens.
XX
PN      WO200110205-A1.
XX
PD      15-FEB-2001.
XX
PF      09-AUG-2000; 2000WO-US21790.
XX
PR      09-AUG-1999; 99US-0370265.
PR      11-JAN-2000; 2000US-0481238.
PR      25-APR-2000; 2000US-0557800.
PR      26-MAY-2000; 2000US-0583231.
PR      30-JUN-2000; 2000US-0608285.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Ford J, Mulero JT, Yeung G;
XX
DR      WPI; 2001-147489/15.
XX
N-PSDB; AAF63383.
XX
PT      Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT      and/or NDPase activity, which are useful in the treatment of
PT      pathological conditions caused by thrombosis (e.g. myocardial
PT      infarction) and inflammatory disorders -
XX
XX
XX      Claim 15; Fig 2; 203pp; English.
XX
XX      This invention relates to polynucleotides encoding human CD39-like
XX      polypeptides with apyrase and/or NDPase activity. The polypeptides having
XX      apyrase, including NDPase, activity are useful for inhibiting platelet
XX      function and can therefore be used in the prophylaxis or treatment of
XX      pathological conditions caused by or involving thrombosis or excessive

```

```

CC      coagulation or excessive platelet aggregation, such as myocardial
CC      infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC      artery thrombosis or intracardiac thrombosis, and conditions associated
CC      with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC      modulating disease states (including platelet aggregation, inflammation
CC      and apoptosis) associated with ADP or other purinergic signalling by
CC      reducing the levels of NDPs. The polypeptides are also useful for
CC      prophylaxis or treatment of inflammation related disorders, such as
CC      disorders involving sepsis or systemic inflammatory response syndrome or
CC      SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC      cytokine overstimulation), autoimmune disorders such as thrombosis,
CC      atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC      cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC      neurological disorders including neurodegenerative diseases, epilepsy,
CC      depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC      disease, and amyotrophic lateral sclerosis; and cancer. The present
CC      sequence represents human CD39 like protein CD39-L4.
XX
SQ      Sequence 428 AA;

```

Query Match 20.6%; Score 491.5; DB 22; Length 428;
 Best Local Similarity 31.3%; Pred. No. 1e-36;
 Matches 134; Conservative 71; Mismatches 166; Indels 57; Gaps 14;

```

Qy      41 LLEKNDPEPTSYAVIPDAGSTGRHVYNFDQNL-DLPEVNELEPYDSVKPGLSSYAN 99
      | :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      38 mcpilvasastlygimfdagstgrlhytftvgkmpgqlpl-legevfdsvkpglsafvdq 96
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      100 PEEAAESTLPIELKEAENVVVSQPTNPVKLGATAGLRILEGNAANITIQVADMISNRS 159
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      97 pkqgaetvglllevakdsiprshkktvpllxatagrlilpekkakallfevkeft-rks 155
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      160 ALNVOSDAVSIIDTQEGSYLWNTINYLGLKLRFTKTGVVDLGGSSVQMTYVNSRNT 219
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      156 pflvpgsvsindgsdeglawtvtvfltgqlhghrgetvqclidlgastqtlfl--pgf 213
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      220 AANAPVPEGEDPYIKLVLOGKKYDLVHVSYLRYGREAFR-----AEIRKV 266
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      214 ektleqtprg---yltsfemfnslyklyhsylgfglkaarlalgaletegtdhtfrs 270
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      267 AGGSANPCILAGFDGAVTYTSGAEKVSAPASGS-NLNOCRKIALAKLVNAPCPYONCTF 325
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      271 a-----clprvlaeawifgvkyvgngqgevegfepeyaevlrvvgklhqpce--- 320
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      326 GGIMWNGGGSGQKNLFLTSSFYI-LSEDVGIFVNNKPNKIRPVDLTKTAACKTNLEDA 384
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      321 -----vgrgsfyafsyddravtdmdldyegkqllkvedferkarevc-dnlenf 369
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      385 KSKYPDLYEKDSVEYVCLDIYVYVTLVDGSLDFQEVYVANELEYDAILVEAAMPIGT 444
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      370 tsgsp-----flcmdlsyitalikdfig---fadstvl-qltkkvmnietgwaiqa 416
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy      445 ALEAISL 452
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      417 tfhlqsl 424

```

```

RESULT  12
AAB72243
ID      AAB72243 standard; Protein: 428 AA.
XX
AC      AAB72243;
XX
DT      14-MAY-2001 (first entry)
XX
DE      Human CD39 like protein CD39-L4 amino acid sequence.
XX
KW      Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW      myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW      cerebral artery thrombosis; platelet aggregation; inflammation;
KW      apoptosis; autoimmune disorder; neurological disorder;
KW      Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.

```


(NRPasee). It is isolated from the human foetal liver-spleen cDNA library b2HASE20W. It is a soluble ATP diphosphoryltransferase (ATPase) and is involved in the hydrolysis of adenosine diphosphate (ADP), the agonist that causes platelet aggregation. CD33-14 protein has 30% and 80% homology to human and murine CD39. It has platelet aggregation inhibition and antithrombotic activity. CD39-14 is used to treat or prevent CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is also used in vitro, to maintain vascular grafts or during extracorporeal circulation, to hydrolyse NRP, as molecular weight markers and as nutritional supplements. It is used to identify therapeutic agents that bind and modulate CD33-14. It is coupled to toxins for targeting drugs CC to tumours or other cells that express CD39-14. CC

SQ Sequence 428 AA:

Query Match	20.48;	Score 486.5;	DB 21;	Length 428;
Best Local Similarity	31.18;	Pred. No. 2.9e-36;		
Matches 133; Conservative	71;	Mismatches 167;	Indels 57.	Gaps 14

[illegible]

RESULT 14
AAB72240
ID AAB72240 standard; Protein; 428 AA.

DT 14-MAY-2001 (first entry)

Human CD39 like protein CD39-L4 variant ACR11 amino acid sequence.

KM human C339-1-like protein; apyrase; NPase; platelet function inhibitor;
KM myocardial infarction; cerebral ischemia; angina; arterial thrombosis
KM cerebral artery thrombosis; platelet aggregation; inflammation;
KM apoptosis; autoimmune disorder; neurological disorder; mutant;
KM Alzheimer's disease; Parkinson's disease; cancer; C339-14.

OS Homo sapiens.

PN W0200110205-A1.

...

PD 15-FEB-2001.

PF 09-AUG-2000; 2000WO-US21790.

PR 09-AUG-1999; 99US-0370265.
PR 11-JAN-2000; 2000US-0481338

PR 26-MAY-2000: 2000US-0583231

XX

XX

XX
XX
EPT. 2001-147400.115

DR N-PSDB; AAF63385.
XX

PT and/or ND Pase activity, which

PT infarction) and infl

PS Claim 19; Fig 6; 203

This invention relates to polynucleotides encoding human CD39-like polypeptides with sequence and/or structure similar to that of

function and can therefore be used
ARbase, including NDBase, active

pathological conditions caused by or involving thrombosis or excessive coagulation or excessive platelet aggregation, such as myocardial infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral artery thrombosis or intracardiac thrombosis, and conditions associated with venous thrombosis. CD39-14 and CD39-12 polypeptides are useful in modulating disease states (including platelet aggregation, inflammation and apoptosis) associated with ADP or other purinergic signalling by reducing the levels of NDPs. The polypeptides are also useful for prophylaxis or treatment of inflammation related disorders, such as disorders involving sepsis or systemic inflammatory response syndrome or SIRS (and associated conditions such as fever, tachycardia, tachypnea, cytokine overstimulation); autoimmune disorders such as thrombosis, atherosclerosis, acute pancreatitis, dermatitis, including psoriasis, cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis, neurological disorders including neurodegenerative diseases, epilepsy, depression, Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis; and cancer. The present sequence represents human CD39-like protein CD39-14 variant ACRII.

Sequence 428 AA;

Query Match	20.48;	Score 486.5;	DB 22;	Length 428;
Best Local Similarity	31.18;	Pred. No. 2.9e-36;		
Matches 133; Conservative	71;	Mismatches 167;	Indels 57;	Gaps 14

41 L L P K N Q E P V T S Y A V I E D A G S T G S R V H Y N F D Q N L - D L P V E N E L E F Y D S V K P G L S S Y A A N 99

Db 38 mcpinvsastlygimfdagstgtrihvyttfvqkmpgqlpl-legevfdsvkpglsafvdq 96

100 PEEAAESLIPLLKEAENVVPSQQPNTPKRLGATAGLRLLEGNAENILQAVRDMLSNRS 155

155

[illegible][illegible][illegible]

214 ekltetqprg---ylstststnstykylytshylygtlkaarlattlgaletegtgdhtfrs 270

201 A00SANPCILHGF DGAI IISGAIEKVSAPASGS -NLNQCKRIALKALKVNAPCPYQNCTF 325

271 a-----clprwleaewifgvgkgygngngegevgfepcyaevlrvvrqklhppee----- 320

QY 326 GGIMNGGGSGOKNLFITSSFFY-LSEYGVFNKRNKIRPVDLKTAKKTLEDA 384
 Db 321 -----vgrgsfyatsyydravtdcmidyekgylkvedferkarevc-dhlfenf 369
 QY 385 KSKYPDLKEDSEYVCLDLVYVYTLVDGFGIDPFOEVTANETLEYQDALVEAMPLCT 444
 Db 370 tsgsp-----flcmldsyitallkdgfgf---fadstvl-qltkkvmnletyqwalga 416
 QY 445 AIEAIISSL 452
 Db 417 tflhlqsl 424

RESULT 15
 AAY44851
 ID AAY44851 standard; Protein; 405 AA.
 AC AAY44851;
 XX
 DF 18-MAY-2000 (first entry)
 XX
 DE Human CD39-L66 protein.
 XX
 KW CD39-L4; human; CD39-L66; apyrase; nucleotide diphosphatase; NDPase;
 KW ATP diphosphohydrolase; ATPDase; adenosine diphosphate; ADP; treatment;
 KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
 KW cerebral ischemia; angina; vascular graft; extracorporeal circulation;
 KW molecular weight marker; nutritional supplement; tumour; prevention;
 KW drug targeting; splice variant.
 OS Homo sapiens.
 XX
 PN WO200004041-A2.
 XX
 PD 27-JAN-2000.
 XX
 PF 16-JUL-1999; 99WO-US16180.
 XX
 PR 16-JUL-1998; 98US-0118205.
 PR 24-JUL-1998; 98US-0122449.
 PR 04-FEB-1999; 99US-0244444.
 PR 19-MAR-1999; 99US-0273447.
 PR 09-JUL-1999; 99US-0350836.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Ford J, Mulero J;
 DR WPI: 2000-182397/16.
 DR N-PSDB: AAZ50358.
 XX
 PT New nucleic acid encoding human CD39-Like protein, useful for treating
 PT and preventing thrombotic disease -
 XX
 PS Claim 15; Page 124-125; 125pp; English.
 XX
 CC The present amino acid sequence is the CD39-L66 protein, a splice
 CC variant of the CD39-L4 protein. It is an apyrase and/or nucleotide
 CC diphosphatase (NDPase), isolated from the human foetal liver-spleen cDNA
 CC library, bZHFLS20W. It is a soluble ATP Diphosphohydrolases (ATPDase) and
 CC is involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
 CC that causes platelet aggregation. CD39-L4 protein has 308 and 808
 CC homology to human and murine CD39. It has platelet aggregation inhibition
 CC and antithrombotic activity. CD39-L4 is used to treat or prevent
 CC thrombosis, myocardial infarction, cerebral ischemia and angina. It is
 CC also used in vitro, to maintain vascular grafts or during extracorporeal
 CC circulation, to hydrolyse NDP, as molecular weight markers and as
 CC nutritional supplements. It is used to identify therapeutic agents that
 CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs
 CC to tumours or other cells that express CD39-L4.
 XX
 SQ Sequence 405 AA;

Query Match 19.8%; Score 472; DB 21; Length 405;
 Best Local Similarity 31.9%; Pred. No. 5,9e-35;
 Matches 128; Conservative 66; Mismatches 151; Indels 56; Gaps 13;

QY 41 LPPKNOEPTSYAVIFDAGSTGSRVHVNFDDNL-DLPEVEDELEFYDSVKPGLSSYAN 99
 Db 38 mcpilnvasstlygimndastgtrihvfyfvgkmpgqpl-legevfdsvkpglsafvq 96
 QY 100 PEEAASLIPDLKEANVVPVSOQPTPVKLGATAGRLLEGNAENIIOAVRDMISNR 159
 Db 97 pkgaetvqgllevakdsiprshwkktpvilkataqrlilpehkaallfevkef-rks 155
 QY 160 ALNVQSDAVSIIDGTOESYLWWTINYLGLKIGKRTKTGVGVDDGGSSVQMTYANSRNT 219
 Db 156 pflvpkgsvsindgsdeglawvtnfltgqlhghrgetvgclldlggastqlfl--pgf 213
 QY 220 AKNAPKVPEDPEPYIKKVLQGGKKYDLYVHSYLRYGREAFR-----AEIFRV 266
 Db 214 ektlqetprg---yltsfemfnstlyklythsyigfikaarlaltgaletegtldhtfrs 270
 QY 267 AGGSANPCIIAGFDGAVTYSGAEYKVSAPASGS-NLNQCRKIALKALKNVAPCPYQNTF 325
 Db 271 a-----clprvleaeewlfgyvkygngqgevgfepcyaevlrvvrgklbqpee---- 320
 QY 326 GGIMNGGGSGOKNLFITSSFFY-LSEYGVFNKRNKIRPVDLKTAKKTLEDA 384
 Db 321 -----vgrgsfyatsyydravtdcmidyekgylkvedferkarevc-dhlfenf 369
 QY 385 KSKYPDLKEDSEYVCLDLVYVYTLVDGFGIDPFOEYTV 425
 Db 370 tsgsp-----flcmldsyitallkdgfgf---fadstv 398

Search completed: June 28, 2002, 04:00:34
 Job time: 6549 sec

THIS PAGE BLANK (USPTO)

```

Db 1 MELLIKITLITLLESMFAITSSOYLGNLTLKRIPL-KQETSSAVVDFDGSGRTH 58
OY 67 YVNFQNDLLVVENLEFYDYKPOLSSYAANPEBAESLIPLEKAENVVPSQOQPT 126
Db 59 VYFNFQNDLLHIGKGVENYKNTIPOLSSYANNPEQAKSLIPLEQADVVDLQPT 118
OY 127 PVKLGATAGLRLLEGNAENIIOAVDMLSNSALNVSDAIIIDGQEGSYLWVTINX 186
Db 119 PVRLGATAGLRLINDASKLIIQSVDMLSNSSTNVQDASIIIDGQEGSYLWVTINX 178
OY 187 LGLGKRTKTVGVYVDLGGGSVQMTYAVSRNTAANAKVPGEDEPYTKLVLQCKYDL 246
Db 179 ALGNLGGKRTKTVGVYVDLGGGSVQMAVSKTAKNAKVDGDDPYIKKVLKQIPYDL 238
OY 247 YVHSLYRGREAFRAEIRFVAGGSANPCLIGEDGATYYSAGEVYSVAPAGSNINQCRK 306
Db 239 YVHSLYHGREASRAEILKLPSPNPCLIGFNGIYYSAGEFKAATYTSGANFNCKKN 298
OY 307 IALKALKVNAPCPYONCTFGGIMNGGSGQKNLFTSSFYLSBDVG-IFVKNPKNAKIR 365
Db 299 TIRKALKINYPCCPYONCTFGGIMNGGSGQKNLFTSSFYLPEDTGMVASTPNTILR 358
OY 366 PVDLTAKTKACKTNLEDAKSKYRDLKESVEYVCLDLYVYTLVDGFGLDPPQEVTV 425
Db 369 PVDLTAKTKACKTNLEDAKSKYRDLKESVEYVCLDLYVYTLVDGFGLDPPQEVTV 418
OY 426 ANEIEYODALVEAAMPILGTAIEATISLTKFPERLWFEI 462
Db 419 GKEIEYODALVEAAMPILGNAVEAISALPKFERLWFEV 455

RESULT 2
APY_SOLUTU STANDARD; PRT; 454 AA.
ID APY_SOLUTU STANDARD; PRT; 454 AA.
AC P80595; Q43164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Aprtase precursor (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine diphosphatase) (ADPase) (ATP-diphosphohydrolyase).
GN ROP1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.
RC TISSUE= tuber;
RX MEDLINE=96158965; PubMed=8579614;
RA Handa M., Guidotti G.;
RT "Purification and cloning of a soluble ATP-diphosphohydrolyase (aprtase) from potato tubers (Solanum tuberosum).";
RL Biochem. Biophys. Res. Commun. 218:916-923(1996).
RN [2]
RP SEQUENCE OF 42-54; 68-95 AND 236-253.
RC STRAIN=CV. DESIREE;
RX MEDLINE=96355615; PubMed=8703025;
RA Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,
RA Kellum A.M., Mancilla M., Valenzuela M.A., Verjovski-Almeida S.;
RT "Partial purification and immunohistochemical localization of ATP diphosphohydrolyase from Schistosoma mansoni. Immunological cross-reactivities with potato aprtase and Toxoplasma gondii nucleoside triphosphate hydrolase.";
RL J. Biol. Chem. 271:22139-22145(1996).
RN [3]
RP FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF NUCLEOSIDE TRI- AND DI-PHOSPHATES.
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -1- COFACTOR: CALCIUM.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE GDAI / CD39 NTPASE FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U58597; AAB02720.1; -
DR EnPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PF01150; GDAI_CD39_1.
DR PROSITE; PS01238; GDAI_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Calcium; Signal.
FT SIGNAL 1
FT CHAIN 31
FT TRANSMEM 426
FT CARBOHYD 151
FT CARBOHYD 262
SQ SEQUENCE 454 AA; 50041 MW; 9D9FEFA31DA2F52F CRC64;

```

Query Match 49.6%; Score 1182.5; DB 1; Length 454;
 Best Local Similarity 51.1%; Pred. No. 1,1e-80;
 Matches 226; Conservative 85; Mismatches 124; Indels 7; Gaps 4;

```

OY 11 ISLMTFVFMILPAISSOYLGNLTLKRIPLKNOEPTVTAVIDAGSTGSRVHYNF 70
Db 9 IFILALFVLPLSLSKVNAQIPLRLHLSHSH---YAVIPDAGSTGSRVHYNFRF 64
OY 71 DONLDLPEVNELEFYDSVPGLSYVAANDEEAESLIPLEKAENVVPSQOQPTVTL 130
Db 65 DEKLGILPIONNIEFYMAHEPGLSSTAEDEPKAANLEPLDGAEGSVYQELQSEPLEL 124
OY 131 GATAGLRLLEGNAENIIOAVDMLSNSALNVSDAIIIDGQEGSYLWVTINYLCK 190
Db 125 GATAGLRLLEGNAENIIOAVDMLSNSALNVSDAIIIDGQEGSYLWVTINYLCK 184
OY 191 LGRFRTKTVGVYVDLGGGSVQMTYAVSRNTAANAKVPGEDEPYTKLVLQCKYDL 250
Db 185 LGRDYSTATITDLDGGGSVQMAVSKTAKNAKVDGDDPYIKKVLKQIPYDL 243
OY 251 YLRGREAFAEIRFVAGGSANPCLIGEDGATYYSAGEVYSVAPAGSNINQCRKIALK 310
Db 244 YLNYGGLARAEIFKASRNENPCALEGCDGYSGVGVYKVKARKKSGSMKRCRLRH 303
OY 311 ALKVNAPCPYONCTFGGIMNGGSGQKNLFTSSFYLSBDVG-IFVKNPKNAKIR 369
Db 304 ALKVNAPCPYONCTFGGIMNGGSGQKNLFTSSFYLSBDVG-IFVKNPKNAKIR 363
OY 370 KTAACKLACKTNLEDAKSKYRDLKESVEYVCLDLYVYTLVDGFGLDPPQEVTVANEI 429
Db 364 LNAKVAQCTNVADKISIRPKTQDR-NIPYLCMDLIYETTLVDGFGLNPKHEITVHDV 422
OY 430 EYODALVEAAMPILGTAIEATISLTKFPERLWFEI 451
Db 423 QYKNYLVGAAMPILGCAIDLVSS 444

RESULT 3
ENP5_HUMAN STANDARD; PRT; 428 AA.
ID ENP5_HUMAN STANDARD; PRT; 428 AA.
AC O75356;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Echonnucleoside triphosphate diphosphohydrolase 5 precursor (EC 3.6.1.6) (NTPdase) (Nucleoside diphosphatase) (CD39 antigen-like 4) (ER-UDPase).
GN ENP5 OR CD39L4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX TISSUE=Leukemia;
 RA MEDLINE=98341119; PubMed=9676430;
 RT Chadwick B.P., Fritschau A.-M.;
 RT "The CD39-like gene family: identification of three new human members
 (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
 the gene family from *Drosophila melanogaster*.";
 RL Genomics 50:357-367(1996).
 CC -|- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
 CC GLYCOPROTEIN FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
 CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
 CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
 CC (BY SIMILARITY).
 CC -|- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
 CC nucleotide + phosphate.
 CC -|- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
 CC -|- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
 CC -|- TISSUE SPECIFICITY: EXPRESSED IN ADULT LIVER, KIDNEY, PROSTATE,
 CC TESTIS AND COLON. MUCH WEAKER EXPRESSION IN OTHER TISSUES.
 CC -|- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 DR EMBL: AF039918; AAC39885.1; -
 DR MIM: 603162; -
 DR InterPro: IPR00407; GDA1_CD39_NTPase.
 DR Pfam: PF01150; GDA1_CD39; 1.
 DR PROSITE: PS01238; GDA1_CD39_NTPASE; FALSE NEG.
 DR Hydrobase: Transmembrane; Glycoprotein; Calcium; Magnesium;
 KM Endoplasmic reticulum; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 428
 FT ECTONUCLEOSIDE TRIPHOSPHATE
 FT DIPHOHYDROLASE 5.
 FT CARBOHYD 42 42
 FT CARBOHYD 232 232
 FT CARBOHYD 368 368
 FT SEQUENCE 428 AA; 47517 MW; 830437A15DE4DD CRC64;
 SQ
 Query Match 20.6%; Score 491.5; DB 1; Length 428;
 Best Local Similarity 31.3%; Pred. No. 3, 1e-29;
 Matches 134; Conservative 71; Mismatches 166; Indels 57; Gaps 14;
 OY 41 LLPKNOEPTVAIVFDAGSTGRVHVNFQDL-DLLEVELEFEYSVKKGLSSYAAAN 99
 DB 38 MCPINVASSTLIGIMEDASTGTIRHVFVQKMPQLPI-LEGEFVDSVKPCLSAFVQ 96
 OY 100 PEAAESLIPLKEAENVVPGQPTPYKLTATGLRLLEGMAANIIQAVRDMLSNS 159
 DB 97 POGAETVGLLEVANDSTPRSHMKTPVLAATAGLRLLPEKAKALLFEVKEIF-RRS 155
 OY 160 ALNVSADAVSILDGNOESYLVWTINYLKGLKRPRTKTVGVVDLAGGSVOMTYAASRMT 219
 DB 156 PLYVRGVSIMDSDEGLIANTVNFILGQLHGRQETVGLDLOGASTQITFL--PGF 213
 OY 220 AKNAKRVPEGEDPYIKLVLGKKYDLYVHSYLRGREAFR-----AEIFRV 266
 DB 214 EKTLEOTPRG---YLTSFEMFNSTYKLYTHSYGLFKAAARLATLGALETEGDTGHTFRS 270
 OY 267 AGSNAKPIAGPDGAYTYSGAEVKYSAPASGS-NLNQKRIATLAKLXNAPPCYNCF 325
 DB 271 A-----CLPRLEAMITGVKTYQYGNOEGEVGEPCTAEYLKRVVRGKLHPDE--- 320
 OY 326 GGIMNGSGSGKNTLFTSSFYV-LSEDVGIYFNKPNKIRPVLDLTAACKTLEDA 384

DB 321 -----VQGSFYAESYYDRAVDMIDYKGLKLVDEPFERKAREVC-DNLNEF 369
 OY 385 KSKYPPDYKDSVEYVCLDYVYTTLLVDFGDPDPOEVYVAIEYQALVNAAPLGT 444
 DB 370 TSGSP-----FLCMLDSTYTLALKDGF--PADSTVL-QLTKRVNIETGMALGA 416
 OY 445 AIEAIISSL 452
 DB 417 TFFHLOSL 424
 RESULT 4
 ENPS_MOUSE
 ID ENPS_MOUSE STANDARD; PRT; 427 AA.
 AC 09NUZ9; 070214;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
 DE (EC 3.6.1.6) (NTPases5) (Nucleoside diphosphatase) (CD39 antigen-like
 DE 4) (ER-NDPase).
 DE ENTPD5 OR CD39L4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX TISSUE=Testis;
 RX MEDLINE=98119025; PubMed=9457681;
 RA Chadwick B.P., Williamson J., Sheer D., Fritschau A.-M.;
 RT "cDNA cloning and chromosomal mapping of a mouse gene with homology to
 RT NTPases.";
 RL Mamm. Genome 9:162-164(1998).
 RN [2]
 RC SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
 RX TISSUE=Liver;
 RX MEDLINE=99298181; PubMed=10369669;
 RA Trombetta E.S., Melenius A.;
 RT "Glycoprotein reglycosylation and nucleotide sugar utilization in the
 RT secretory pathway: identification of a nucleoside diphosphatase in the
 RL endoplasmic reticulum.";
 RL EMBO J. 18:3282-3292(1999).
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arkawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Radota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fieleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuenl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli I., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carinici P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -|- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
 CC GLYCOPROTEIN FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
 CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
 CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE
 CC PYROPHOSPHATE.
 CC -|- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a


```

OY 127 PVKLGATAGLRLEGNAENIIOAVRDMLSNRSALNVOSDAVSLDGTQEGSLWNTINX 186
DB 124 PVLKATAGLRRLPEKKAQALLLEVEIR-KMSPLVPDSDSINDSGEGLIANTVNF 182
OY 187 LUGKLRKRTKTVGVVDLGGSSVOMTYAVS-RNTAKNAKRVBEGEDPYIKLVLOGKKYD 245
DB 183 LFGOLHGHSEETMGTLDLGGASTQTLFPGSKTLEQTRP-----DYLTFEMENSTPK 236
OY 246 LVHSHYLRGRGAFR-----AEIKVAGGSANPCILAGFDATYYSCAEYV 292
DB 237 LKTHSYLGGLKAARATLGALETGCTDCHTRSA-----CLPRMLEAMFEGGKXOY 290
OY 293 SAPASGS-NLNCCKRIATKAL--KVNAPCYONCTFGGIMWGGSGQNLPLTSSFYVL 349
DB 291 GGNQSGEMGEPCYAEVLRVVOGKHLHPEIRGSSP-----YAFSTYIDR 335
OY 350 SEDVGIFFVNRPAKIRPVDLKTAAKLACTNLEDAKSKYPDYEKDSVEYVCLDLYVYT 409
DB 336 AADTHLIDYKGGVLAKEVDFEKARREV-C-DNLESFTSGSP-----FLCMDLSYIRA 385
OY 410 LLYDGFGLDPFOEVYVANEIYQDALVEAPPLG 443
DB 386 LKDGSG--FADGTL-QLTKRVNIETGMPDG 415

RESULT 6
ENP6_HUMAN STANDARD: PRT; 484 AA.
ID ENP6_HUMAN
AC 075354; Q9JUD1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ectonucleoside triphosphate diaphosphohydrolase 6 (EC 3.6.1.6)
DE (NTPDase6) (CD39 antigen-like 2).
GN ENP6 OR CD39L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=keralinocytes;
RX MEDLINE=98341119; PubMed=9676430;
RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clump M., Clark L.N., Clark S.Y., Clee C.M.,
RA Clagg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,
RA Coulson A., Coville G.J., Deedman R., Dhani P., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Williams R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Leharasainio M.H., Leversha M., Lloyd G., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McCormack L.J., McIay R., McMurtry A.A.,
RA Mlynar S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,
RA Swann M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

```

```

RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -I- FUNCTION: MIGHT SUPPORT GLYCOSYLATION REACTIONS IN THE GOLGI
CC APPARATUS AND, WHEN RELEASED FROM CELLS, MIGHT CATALYZE THE
CC HYDROLYSIS OF EXTRACELLULAR NUCLEOTIDES. HYDROLYZES PREFERENTIALLY
CC NUCLEOSIDE 5'-DIPHOSPHATES. NUCLEOSIDE 5'-TRIPHOSPHATES ARE
CC HYDROLYZED ONLY TO A MINOR EXTENT. THERE IS NO HYDROLYSIS OF
CC NUCLEOSIDE 5'-MONOPHOSPHATES. THE ORDER OF ACTIVITY WITH DIFFERENT
CC SUBSTRATES IS GDP > IDP > UDP = CDP >> ADP (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -I- CORRELATION: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC OCCURS IN A SOLUBLE EXTRACELLULAR FORM (BY SIMILARITY).
CC -I- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES.
CC -I- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC EMBL; AF039916; AAC39883.1; -.
CC EMBL; AL035252; CAB41571.1; -.
CC MIM; 603160; -.
CC InterPro: IPR000407; GDAL_CD39_NTPase.
CC Pfam: PF01150; GDAL_CD39_1; NTPase; FALSE_NEG.
CC PROSITE: PS01238; GDAL_CD39_NTPase; FALSE_NEG.
CC KEGG: Hydrolyase; Transmembrane; Glycoprotein; Calcium; Magnesium;
CC KW SIGNAL-anchor; Golgi stack.
CC FT DOMAIN 1 39 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 40 60 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC FT (POTENTIAL).
CC FT DOMAIN 61 484 LUMENAL (POTENTIAL).
CC FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 138 138 V -> L (IN REF. 2).
CC FT CONFLICT 202 202 E -> K (IN REF. 2).
CC FT SEQUENCE 484 AA; 53233 MW; 27334B290B8064C CRC64;
SO

Query Match 19.4%; Score 463; DB 1; Length 484;
Best Local Similarity 31.3%; Pred. No. 4.9e-27;
Matches 131; Conservative 70; Mismatches 156; Indels 62; Gaps 13;

```



```

CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19560; AAA34656.1; -
DR EMBL; U18779; AAB65000.1; -
DR PIR; S30837; S30837.
DR PIR; A40732; A40732.
DR SGD; S0000768; GDAL.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01238; GDAL_CD39_NTPase; 1.
KW Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT LUMENAL (POTENTIAL).
FT CARBOHYD 41 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 518 AA; 56821 MW; 9A61182D5ED22ADF CMC64;

Query Match 18.6%; Score 442.5; DB 1; Length 518;
Best Local Similarity 29.1%; Pred. No. 1.8e-25;
Matches 153; Conservative 70; Mismatches 178; Indels 125; Gaps 20;

QY 17 VPMALPAISSQY-----LGNNTIMRKRT-----LLPKNDP----- 48
DB 17 VIMLILKRTSSIGPPSARTVTPNASIPKTPEDISILPVNDEPYLDQSKTEQNPETIA 76
QY 49 -----VTS-----YVIFDAGSGSRVHYNFD-----QNDLIPVENELFFYDSV 89
DB 77 DAVKSQTSQTSSEHKYIYIMDAGSGSRVHYNFDVCTSPPTLD-----EKFDWL 128
QY 90 KPGLSYVANDPEEAESLIPLEKNEVNVVSOQPTPVKLGATAGLRLEGNNAENTIQ 149
DB 129 EPGLSFPTDSVGAANSIDPLKLVAMNYVPIKARCTPVAVATAGLRLLGDAKSKILTS 188
QY 150 AVRODLSNRSLN-VQSAVSTLSTOGSGSYMTNTINILGKGRFRK--TVGVVDLGG 206
DB 189 AVRODLERDYPPEVVEGSGVSGIMSGDEGVFAMITNTNLLGNIGANGPKLPAAVFDLGG 248
QY 207 GSVQMTYAVSNRTAKNAKRVPEGEDPYIKKLVLQKKYDLVHSYLRYGREAFRAIFKV 266
DB 249 GSTQIVF---EPTFPINERKMDGEHKF---DLKFGDENTLIVQSHLGTGLKGRKRVASV 303
QY 267 AGGSA--NPCILAFGDGYVTVSGAEYKVASAPASGSNLN----- 302
DB 304 LVENALKDKGKILKG-DNRTKT-----HQLSSPCLPKPVNATNEKYTLSEKETYTIDFIPGD 357
QY 303 -----QCRKIALKALKVNAPCYONCTGFGITNGG---GGSQKRLFLTSSYYTISEDVG 354
DB 358 EPGSAQCFPLDELILNKDAQCPSPCSFNGVQPSLVRTFKESNDIYIFSYFYDTRPLG 417
QY 355 IFVNPKNKIRPV-----DLTKAALACKTINLEDAKSKYP---DLYEKDSVEYVCLDIY 405
DB 418 M-----PLSFTLMEINDLARIYCKGE-ETWNSVFSGAGSLDELDESISHRCLDLS 466
QY 406 VYVITLLVDGFGIDPFQEVTVANEIEYQDALVEAMPICGTAIEAISS 451
DB 467 FOVSILHTGYDIPLORELRTGKRI---ANKETIGCLGASLPLAKA 508

RESULT 9
ENPL_RAT STANDARD; PRT; 511 AA.
ID ENPL_RAT

```

```

AC P97687;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NMPDase), (ecto-ATP diphosphohydrolase) (ATPDase) (lymphoid cell
DE activation antigen) (CD39 antigen) (Ecto-apyrase).
GN ENMPD1 OR CD39.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain, and Hippocampus;
RX MEDLINE=96215267; Pubmed=8626624;
RA Wang T.-F., Rosenberg P.A., Guidotti G.;
RT "Characterization of brain ecto-apyrase: evidence for only one ecto-
RT apyrase (CD39) gene.";
RL Brain Res. Mol. Brain Res. 47:295-302(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Cochlea;
RX MEDLINE=20050856; Pubmed=10581401;
RA Vlakovic S.M., Housley G.D., Greenwood D., Thorne P.R.;
RT "Evidence for alternative splicing of ecto-ATPase associated with
RT termination of putative transmission.";
RL Brain Res. Mol. Brain Res. 73:85-92(1999).
RN [3]
RP SEQUENCE OF 432-511 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=98031057; Pubmed=9364474;
RA Kegel B., Braun N., Heine P., Maliszewski C.R., Zimmermann H.;
RT "An ecto-ATPase and an ecto-ATP diphosphohydrolase are expressed in
RT rat brain.";
RL Neuropharmacology 36:1189-1200(1997).
CC -1- FUNCTION: IN THE NERVOUS SYSTEM, COULD HYDROLYSE ATP AND OTHER
CC NUCLEOTIDES TO REGULATE PURINERGIC NEUOTRANSMISSION. COULD ALSO
CC BE IMPLICATED IN THE PREVENTION OF PLATELET AGGREGATION.
CC HYDROLYZES ATP AND ADP EQUALLY WELL.
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O -> AMP + 2 phosphate.
CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN PRIMARY NEURONS AND ASTROCYTES,
CC KIDNEY, LIVER, MUSCLE, THYMUS, LUNG AND SPLEEN.
CC -1- PIM: N-GALACTOSYLATED.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U81295; AAC53195.1; -
DR EMBL; Y15685; CAAT5730.1; -
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01238; GDAL_CD39_NTPase; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 37 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 38 478 POTENTIAL.
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 511 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CONFLICT 332 332 F -> L (IN REF. 2)
 SQ SEQUENCE 511 AA; 57408 MW; 4CC123D3B644C193 CMC64;

Query Match 17.8%; Score 424; DB 1; Length 511;
 Best Local Similarity 28.1%; Pred. No. 4.2e-24;
 Matches 132; Conservative 70; Mismatches 163; Indels 104; Gaps 22;

OY 42 LPKNGEPTSVAVIPDASTGSRHVVYFNDQMLLPVENE-----LEFYSVKPEL 93
 DB 43 LPEN-----VKIGVLDAGSSHTNLYIKRW-----PAKENDGTGVQVLLLECCVKGPGI 91
 OY 94 SYAANPEEAESTLPLKEAENVVPSOQDNTPYKLGATAGLRLL-----EGNAENITQA 150
 DB 92 SKYAAKTDEIAIYLAECKMKSTERIPASKQOTYVYLGATIGMRLRMESKQSDAVIAA 151
 OY 151 VRDMISNSALNVOSD--AVSILDGTGSGYLWYNTLGLKLGKRFK----- 197
 DB 152 V-----SRSLKSPDFEGAKITGQEGAGYGITINLLG-----RFTQEGSWINFTSDS 202
 OY 198 -----TVGVVDGGGSGVOMTVAVSNRTAKNAKVEPEGDPYIKLVLOGKTYDVHSTLR 253
 DB 203 OKQATFGALDGGSSTOYTF--VPLNQGLEAPETS-----LQRLYGTDTYVYTHSFLC 254
 OY 254 YGREAF-----RAEIKYVAGGS--ANPCILAGFDGATYYSGA-----EYK 291
 DB 255 YGKQDALMOKLAODIQYSSGGILMDPCFYPKYKVVAVSELYGPGCKRKREKILPENOFO 314
 OY 292 VSAASGSNINOCCKIKALKAVAPCPYONCTFFGIGW-----NGGGSGGKNLELSSFY 347
 DB 315 VQGT-----GDEYQCHQSILKEFN--NSHCPYSCAFNGVFLPLDGSFGA-----FSAFY 363
 OY 348 YLSEDVGIFVKKPAKIRPYDLKT--AAKLCKNTLEDAKSKYRPLYSKDSVEYVCIDLTV 406
 DB 364 FVMD-----FFKKMANDSVSOEKEITKNCCKRWEVKASTPYKVKYLSER--CESGTY 419
 OY 407 VYTLVGVFGIDPEQVYVANELEYODALVE--AAAPLGAIPAISLSP 453
 DB 420 ILTLLDGGYNTGT-----TSMDQIHFMGKIKDSNAGWTLGYMLNLJTNMP 464

RESULT 10
 ENPL_HUMAN STANDARD; PRT; 510 AA.
 AC P4961; Q9U0Q9; Q9Y3Q9;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
 DE (NTPase) (Ecto-ATP diphosphohydrolase) (ATPase) (Lymphoid cell
 DE activation antigen) (CD39 antigen) (Ecto-apyrase).
 GN ECTP1 OR CD39.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM VASCULAR).
 RX MEDLINE=95015846; PubMed=7930580;
 RA Maliszewski C.R., Delaespere G.J.T., Schoenborn M.A., Amlltge R.J.,
 RA Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
 RA Birks C., Albert A., Friend D., Gimpel S.D., Geyle R.B. III;
 RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
 RT structural characterization."
 RL J. Immunol. 153:3574-3583(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM VASCULAR).
 RX TISSUE=umbilical vein;
 RA MEDLINE=97149443; PubMed=8996251;
 RA Robson S.C., Kaczmarek E., Siegel J.B., Candinas D., Kozlak K.,
 RA Millan M., Hancock W.W., Bach F.H.;

RT "Loss of ATP diphosphohydrolase activity with endothelial cell
 RT activation.";
 RT J. Exp. Med. 185:153-163(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS PLACENTAL I AND II).
 RC TISSUE=Placenta;
 RX MEDLINE=99332082; PubMed=10405171;
 RA Matsumoto M., Sakurai Y., Kokubo T., Yagi H., Makita K., Matsui T.,
 RA Titani K., Fujimura Y., Narita N.;
 RT "The cDNA cloning of human placental ecto-ATP diphosphohydrolases I
 RT and II."
 RL FEBS Lett. 453:335-340(1999).
 RN [4]
 RP SEQUENCE OF 240-242; 162-166; 185-189; 210-223; 243-249 AND 288-297.
 RC TISSUE=Placenta;
 RX MEDLINE=9606723; PubMed=8529670;
 RA Christoforidis S., Papamiraki T., Galaris D., Kellner R., Tsolas O.;
 RT "Purification and properties of human placental ATP
 RT diphosphohydrolase."
 RL Eur. J. Biochem. 234:66-74(1995).
 RN [5]
 RP SEQUENCE OF 1-30; 49-65; 94-97; 88-93; 193-223; 245-265; 382-385 AND
 RP 399-405 (ISOFORM PLACENTAL I).
 RC TISSUE=Placenta;
 RX MEDLINE=99062444; PubMed=9846014;
 RA Makita K., Shimoyama T., Sakurai Y., Yagi H., Matsumoto M., Narita N.,
 RA Sakamoto Y., Saito S., Ikeda Y., Suzuki M., Titani K., Fujimura Y.;
 RT Placental ecto-ATP diphosphohydrolase: its structural feature
 RT distinct from CD39, localization and inhibition on shear-induced
 RT platelet aggregation.";
 RL Int. J. Hematol. 68:297-310(1998).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=97115858; PubMed=8955160;
 RA Kaczmarek E., Kozlak K., Seviigny J., Siegel J.B., Anrather J.,
 RA Beaudoin A.R., Bach F.H., Robson S.C.;
 RT "Identification and characterization of CD39/vascular ATP
 RT diphosphohydrolase.";
 RL J. Biol. Chem. 271:33116-33122(1996).
 CC -1- FUNCTION: IN THE NERVOUS SYSTEM, COULD HYDROLYSE ATP AND OTHER
 CC NUCLEOTIDES TO REGULATE PURINERGIC NEUROTRANSMISSION. COULD ALSO
 CC BE IMPLICATED IN THE PREVENTION OF PLATELET AGGREGATION.
 CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
 CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
 CC -1- SUBUNIT: HOMODIMER. DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; VASCULAR (SHOWN HERE), PLACENTAL
 CC I AND PLACENTAL II; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY ON ACTIVATED LYMPHOID
 CC CELLS. ALSO EXPRESSED IN ENDOTHELIAL TISSUES. THE VASCULAR ISOFORM
 CC AND THE PLACENTAL ISOFORM II ARE PRESENT IN BOTH PLACENTA AND
 CC UMBILICAL VEIN, WHEREAS PLACENTAL ISOFORM I IS PRESENT IN PLACENTA
 CC ONLY.
 CC -1- PTR: THE N-TERMINUS IS BLOCKED.
 CC -1- MISCELLANEOUS: OPTIMAL PH IS 7.0-7.5 WITH ATP AS SUBSTRATE AND
 CC 7.5-8.0 WITH ADP.
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD39 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd39.htm".
 CC -----
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S73813; AAB32152.1; -
 CC EMBL: U87967; AAB47572.1; -
 CC EMBL: AJ133133; CAB41886.1; -
 CC EMBL: AJ133134; CAB41887.1; -

DR MIM: 601752; -
 DR InterPro: IPR000407; GDAL_CD39_NTPase.
 DR Pfam: PF01150; GDAL_CD39: 1.
 DR PROSITE: PS01238; GDAL_CD39_NTPase: 1.
 KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium;
 KM Alternative splicing.
 FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 38 478 POTENTIAL.
 FT TRANSMEM 479 499 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 500 510 POTENTIAL.
 FT CARBOHYD 73 73 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1 4 MEDT -> MKGTGKDLTSQ (IN ISOFORM PLACENTAL II AND ISOFORM PLACENTAL II).
 FT VARSPLIC 272 299 VASNEILDPCFHPGKRVVNSDLYKT -> ASITGSRPA
 FT PETSAPAPTSCEFLFQIQ (IN ISOFORM PLACENTAL II).
 FT VARSPLIC 300 510 MISSING (IN ISOFORM PLACENTAL II).
 FT CONFLICT 57 58 SS -> G (IN REF. 5).
 FT CONFLICT 162 162 D -> K (IN REF. 4).
 FT CONFLICT 208 208 T -> TGET (IN REF. 5).
 FT CONFLICT 248 248 V -> Y (IN REF. 5).
 SQ SEQUENCE 510 AA; 57964 MW; BAD87D249649159 CRC64;

Query Match 17.7%; Score 421.5; DB 1; Length 510;
 Best Local Similarity 27.4%; Pred. No. 6,5e-24;
 Matches 126; Conservative 73; Mismatches 174; Indels 87; Gaps 19;

QY 42 LKRNQPTSYAVITDAGSGSRVHYVNDQNLDPVNE-----LEPTDYKPEL 93
 DB 43 LLEN-----VKYGIADVDSHTSLYIKV-----PAKENDTGVAHVGECAVKGPGI 91
 QY 94 SSYANPEEASLSPLPLKEAVNVVVSQPNTPYKLGATAGRLLEGNA---AENITLQA 150
 DB 92 SFYQKVNFEIGIYLDCKMERAREVTPRSOHOETPYILGATAGRLRMESEELADRVLV 151
 QY 151 VEDMLSNRSALNVQSDAVSLDGTQGSYLWYINYLGLKLR-----TKT 198
 DB 152 VERSLSN---YPEFGAGIITGEGGAGWITINVLKFKSGKRWESIVPEYTNQET 208
 QY 199 VGVNVLGGSVQMTAVSNNTAKNAKVPGEDEPYTKKLVLOGKKYDLYVSHYLRGREA 258
 DB 209 FGALDLGGASTQVTPVPOQITLES-----PDN---ALQRLYKGDVNVYTHSLCYGKDO 260
 QY 259 F-----RAEIFKVAAGS--ANPCILAGFDGAVYISG-----AEKVVSAPASG-----S 299
 DB 261 AIMOKLAKIOYASNEILDPCHFHQYKRVVNSDLYKTPCKTKRREMTLPFOQFETIGIG 320
 QY 300 NLNOCRKILKALKVNAFCPYONCTFGGIM---NGGGSGGKNLFLTSSFYLSBDVGI 355
 DB 321 NVOQHQSTLELFTNSY-CFYSQCAFNGFLPPLQGDRA-----FAFTFVAK----- 368
 QY 356 FVNKNPAKIRPVDLTAARLACKTNLEDAKSKYPDLREKDEYEVCLDLYVYVYTLAVD- 414
 DB 369 FLNTLSEKVSQEKVEMMKFCAQPMEEIKTSYAGVKEKYLEY-CFSGTYILSLLDQY 427
 QY 415 -FGLDPEQEVYANLEIYODALVEAAMPGLTAIEAIISSLP 453
 DB 428 HFTADSWEHIHFTIGKIQGSD---AGWTLGMYMLNTNMP 463

RESULT 11
 ENPI_MOUSE STANDARD; PRT; 510 AA.
 AC P55772;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
 DE (NTPdase1) (Ecto-ATP diphosphohydrolase) (ATPase) (lymphoid cell
 DE activation antigen) (CD39 antigen) (ecto-apyrase).
 GN ENPDL OR CD39.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95015946; PubMed=7930580;
 RA Maliszewski C.R., Delaespense G.J.T., Schoenborn M.A., Arnltage R.J.,
 RA Fawlow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
 RA Barks C., Albert A., Friend D., Gimpel S.D., Gayle R.B. III:
 RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
 RT structural characterization".
 RL J. Immunol. 153:3574-3583(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=96399871; PubMed=9730622;
 RA Schoenborn M.A., Jenkins N.A., Copeland N.G., Gilbert D.J.,
 RA Gayle R.B. III, Maliszewski C.R.;
 RT "Gene structure and chromosome location of mouse Cd39 coding for an
 RT ecto-apyrase".
 RL Cytogenet. Cell Genet. 81:287-289(1998).
 CC -I- FUNCTION: IN THE NERVOUS SYSTEM, COULD HYDROLYSE ATP AND OTHER
 CC NUCLEOTIDES TO REGULATE PURINERGIC NEUROTRANSMISSION. COULD ALSO
 CC BE IMPLICATED IN THE PREVENTION OF PLATELET AGGREGATION.
 CC HYDROLYZES ATP AND ADP EQUALLY WELL.
 CC -I- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
 CC -I- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
 CC -I- SUBUNIT: HOMODIMER. DISULFIDE-LINKED (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -I- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF037366; AAB92259.1; -
 DR EMBL: AF041818; AAC83203.1; -
 DR EMBL: AF041813; AAC83203.1; JOINED.
 DR EMBL: AF041814; AAC83203.1; JOINED.
 DR EMBL: AF041815; AAC83203.1; JOINED.
 DR EMBL: AF041816; AAC83203.1; JOINED.
 DR EMBL: AF041817; AAC83203.1; JOINED.
 DR MGI: 102805; Entpdl.
 DR InterPro: IPR000407; GDAL_CD39_NTPase.
 DR Pfam: PF01150; GDAL_CD39: 1.
 DR PROSITE: PS01238; GDAL_CD39_NTPase: 1.
 KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium;
 FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 38 478 POTENTIAL.
 FT TRANSMEM 479 499 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 500 510 POTENTIAL.
 FT CARBOHYD 73 73 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 428 428 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 510 AA; 57205 MW; 8E5A6113D2E13930 CRC64;

Query Match 17.2%; Score 410.5; DB 1; Length 510;

Best Local Similarity 26.5%; Pred. No. 4.3e-23;
Matches 124; Conservative 80; Mismatches 161; Indels 103; Gaps 22;

```

OY 42 LKNEPVSATVDFDAGSGSVHYNDQMLDLPVENE-----LEFYDSVKGGL 93
DB 43 LLEN-----VKIVDADGSSHTNLYTKM-----PAEKEDTGVVOOLECQVKGGI 91
OY 94 SSYAANPEEAESLPLKEAENVVPSQOPNTPVKLGATAGLRLL-----EGNAENILQA 150
DB 92 SKYAKTDEIGAVLAECMELSTELPTKSHQTPYVLCATATAMRLLRMESEGSADVLAA 151
OY 151 VDMMSNSALNVQSDAVSILDGTOEGSYLWVTINYLKGLKREPT----- 196
DB 152 VSTSE-----KSYPPDFOGAKIITGDEGAYGWTINYLGL-----RPTQOSMLSLISDSQK 204
OY 197 -KTGVVDLGGGVSOMTAVSNRTAKNPKPEGEDPIKVLQGGKYDLYHSYLRKG 255
DB 205 QETFPALDLGGASTOITVPONSTIES-----PENS-----LQFRLYGEDYTYVTHSFCLYG 256
OY 256 REAF-----RALEFKVAGGSA--NPCIIAGFDGAYTYSGA-----EYKVS 293
DB 257 DQALMOKRLADIQVSSGVLKPCFENFGREKVVNVSELYGTPCKRPEKLPFDQRIQ 316
OY 294 APASGNNOCRTALKALKNAECPPYONCTFFGIGW-----NGGGSGGKNLFTSSFYEL 349
DB 317 GT-----GDEQGHQSTLELFPN--NSHCPYSQCAFNGVFLPPLHSGFGA-----PSAFYFV 365
OY 350 SEDVGIENVK--PNAKIRPVDLKTAALKACTNLEDAKSKYPDLYEKDSVEYVCLDLYV 407
DB 366 MD-----PEFKYAKNSVSIQEMKELTETKNECKSWEERTKTSIPSKYKELTSEY-CFSGAYI 420
OY 408 YTLVLDGGL--DPPQEVTVANEIEYODALVEAAMPJGTAIEAIISSIP 453
DB 421 LSLT-OGVNFDTSSMEDIHFMGIKIDSN---AGWTLGMYMLNTLMIP 463

RESULT 12
ENPL_BOVIN
ID ENPL_BOVIN STANDARD; PRT; 513 AA.
AC 018936;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE activation antigen) (CD39 antigen) (Ecto-apyrase).
GN ENTPD1 OR CD39.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Aortic endothelium;
RA Chang A.S., Garcia R.L., Chang S.M., Schilling W.P.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 97-103; 123-133; 136-140; 145-168 AND 459-471.
RC TISSUE-Aorta;
RA MEDLINE=97115858; PubMed=8955160;
RA Kaczmarek E., Kozlak K., Seviary J., Siegel J.B., Anrather J.,
RA Beaudoin A.R., Bach F.H., Robson S.C.;
RT Identification and characterization of CD39/vascular ATP
RT diphosphohydrolase";
RL J. Biol. Chem. 271:3116-3122(1996).
CC -1- FUNCTION: IN THE NERVOUS SYSTEM, COULD HYDROLYSE ATP AND OTHER
CC NUCLEOTIDES TO REGULATE PURINERGIC NEUROTRANSMISSION. COULD ALSO
CC BE IMPLICATED IN THE PREVENTION OF PLATELET AGGREGATION.
CC HYDROLYSES ATP AND ADP EQUALLY WELL.
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF005940; AAB62382.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
DR PROSITE; PS01238; GDA1_CD39_NTPase; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT DOMAIN 1 16
FT TRANSMEM 17 37
FT DOMAIN 38 481
FT TRANSMEM 482 502
FT DOMAIN 503 513
FT CARBOHYD 73 73
FT CARBOHYD 227 227
FT CARBOHYD 245 245
FT CARBOHYD 307 307
FT CARBOHYD 336 336
FT CARBOHYD 373 373
FT CARBOHYD 460 460
FT CONFLICT 97 97
FT CONFLICT 101 103
FT CONFLICT 464 464
SQ SEQUENCE 513 AA; 58113 MW; 20FE98B27B6D2FE96 CRC64;

Query Match 17.0%; Score 406; DB 1; Length 513;
Best Local Similarity 26.4%; Pred. No. 9.3e-23;
Matches 134; Conservative 90; Mismatches 168; Indels 116; Gaps 25;

OY 10 LISIMTFVFMIMPAISSQYLGNNILMNRKILLPKNKEPTSVAYIPDAGSGSVHYNN 69
DB 17 ILSILGFSCLII-AVALALALG-----LIQNKALPEN---VFGLVLDGSSHTSLYTR 66
OY 70 PQDNLDLPEVENE-----LEFYDSVKGGLSSYAANPEEAESLPLPLKEAENVVPS 121
DB 67 W-----PAEKENDGVYTOLEESVVGKPGISGFAKVVNEINVLVTOMERAQVITSI 119
OY 122 QQPNTPVKLGATAGLRLL-----EGNAENILQVNRMLSNRSLANVQSDAVSILDGTOGCS 178
DB 120 QHMETPVYVIGATAGMKLMLMKENKQMAADKILAV---ASSISPPDFOGARLIISQDEGA 176
OY 179 YLWVTINYLGLKGRFT-----KTGVVDLGGGVSOMTAVSNRTAKNAPKV 226
DB 177 YGMITVNVYLKGFQTKLSMFNLKPSKDDQIETRYGALDLGASTQITF-VPNQETTESP- 233
OY 227 PEGEDPIKVLQGGKYDLYHSYLRIGREAFRAEIFKVAGGSA-----NPCIIAGF 279
DB 234 --NNMLYFR--LYGKNYSVYTHSFCLCYGKD--QALQKLAGLOGTGNGIHEPC-----F 282
OY 280 DGAYTYSAGYKVSAPASG-----SNLNOGRTALKALKAYNA 316
DB 283 HSRYM---RKIMASVLNNEGFCRKHEILNSRPPLVDITIRAGANGFORKOSITIQLFNTSY 339
OY 317 PCPYONCTFFGIGW-----NGGGSGGKNLFTSSFYLSDEVGIFVKNNAKIRPVDLKTA 372
DB 340 -CPYSSCSFNGVFLPPLHGORFGA-----PSAFYVME---FLNLTSESVSVEQIT- 386
OY 373 AKLA--CTNLEDAKSKYPDLYEKDSVEYVCLDLYVTTLYVDG--RGLDPPQEVTVANE 428
DB 387 EKLIEFCAQRMEEVQKNGEVEKELTSEY-CFSGTYILVLLINGVHTFAESWKNHIMNMK 445
OY 429 IEYODALVEAAMPJGTAIEAIISSLPKFE 456
DB 446 VASSTD-----VGWTLGMYMLNTLNKIPAE 469
```


RT	*Evidence for alternative splicing of ecto-ATPase associated with	
RT	termination of putinergic transmission.?	
RL	Brain Res. Mol. Brain Res. 73:85-92(1999).	
CC	-1- FUNCTION: IN THE NERVOUS SYSTEM, COULD HYDROLYZE ATP AND OTHER	
CC	NUCLEOTIDES TO REGULATE PUTINERGIC NEUROTRANSMISSION. HYDROLYZES	
CC	ADP ONLY TO A MARGINAL EXTENT.	
CC	-1- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + phosphate.	
CC	-1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).	
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE	
CC	PRODUCED BY ALTERNATIVE SPLICING.	
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART, VAS DEFERENS,	
CC	KIDNEY, SKELETAL MUSCLE, THYMUS, LUNG AND SPLEEN. WEAK EXPRESSION	
CC	IN LIVER.	
CC	-1- INDUCTION: BY FSH IN SERTOLI CELLS BUT NOT IN PERITUBULAR CELLS;	
CC	BY CAMP IN BOTH TYPE OF CELLS.	
CC	-1- PTM: HAS PROBABLY DISULFIDE BONDS.	
CC	-1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.	
CC	-----	
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb.ch/announce/	
CC	or send an email to license@isb.slb.ch).	
CC	-----	
DR	EMBL; Y11835; CAA72533.1; -	
DR	EMBL; AF276940; AAF87740.1; -	
DR	EMBL; AF129103; AAD42303.1; -	
DR	InterPro; IPR000407; GDA1_CD39_NTPase.	
DR	Pfam; PF011150; GDA1_CD39; 1.	
DR	PROSITE; PS01238; GDA1_CD39_NTPASE; 1.	
KW	Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;	
KW	Alternative splicing.	
FT	DOMAIN	1 4
FT	TRANSMEM	5 25
FT	DOMAIN	26 462
FT	TRANSMEM	463 483
FT	DOMAIN	484 495
FT	CARBOHYD	64 64
FT	CARBOHYD	129 129
FT	CARBOHYD	294 294
FT	CARBOHYD	306 306
FT	CARBOHYD	319 319
FT	CARBOHYD	378 378
FT	CARBOHYD	443 443
FT	VANSPPLIC	466 495
FT	CYTOPLASMIC (POTENTIAL).	
FT	POTENTIAL.	
FT	EXTRACELLULAR (POTENTIAL).	
FT	POTENTIAL.	
FT	CYTOPLASMIC (POTENTIAL).	
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	VRSKSGCAL -> DVRSQPTVQSEVHSEMDFCDDLOGPN	
FT	FLSGPLRQAPPEPGLMESVPLVKTIVYIDFS (IN	
FT	ISOFORM 2).	
FT	T -> A (IN REF. 2).	
FT	PF -> IL (IN REF. 2).	
FT	A -> T (IN REF. 2).	
FT	L -> F (IN REF. 2).	
FT	CONFLICT 444 444	
FT	SEQUENCE 495 AA; 54389 MM; 237B999FLBEB8E00 CRC64;	
QY	Query Match 15.9%; Score 378; DB 1; Length 495;	
QY	Best Local Similarity 26.3%; Pred. No. 1,1e-20;	
QY	Matches 119; Conservative 83; Mismatches 170; Indels 80; Gaps 20;	
Db	46 QEPYTSYAVIFDASTGSRVAYNF--DQNDILPVENELEFYDSVKRGLSSVYANPEEA 103	
Db	34 EPPALKRGIVLDAAGSSHTSMFYKMPADKEND-GGIYQHSSCCVQGGISSVYANDSPA 92	
QY	104 AESLIPLKLEAVENVPVSOQNPVKGATAGNR---LLBENAMENLQAVRWDLNSN--- 157	
Db	93 GQSLVRCLEQALRVPYPRDRASSTPLVIGATRGAMPENULTSEAPARVALEAVOTRTQYPP 152	
QY	158 --REALNVQSDAVSILDTQGSYLWYINVL-----GLTK--RFTK-TVGVVDLG 206	
Db	153 DFRGA-----RLISGQDEGVGCVWTANILENFIKYGWGVGWINPRGGTIGAMLDGG 204	

[illegible]

Search completed: June 28, 2002, 04:04:43
Job time: 722 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 00:42:50 ; Search time 254.74 Seconds
(without alignments)
10035.666 Million cell updates/sec

```

Title:      US-09-657-631-8
Perfect score: 1489
Sequence:   1 aagtgcctctctctctgtag.....cgacatcttctctcctgctt 1489

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 34728722
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneset.032802.*

1:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA2001.DAT.*
23:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA2002.DAT.*
24:	/SIDS1/gcgdata/geneseq/genseqn-emb1/NA2002.DAT.*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1479.6	99.4	1489	20	AAK08529	NBP46 (root lectin
2	869.2	58.4	1458	20	AAK08528	NBP46 (root lectin
3	839.8	56.4	1661	21	AAA96064	Ecto-phosphatase c
4	835	56.1	1667	22	AAF85681	Pea blight resistar
5	833.4	56.0	1697	22	AAF85680	Pea blight resistar
6	830.4	55.8	1368	22	AAF85679	Pea blight resistar
7	824.6	55.4	1643	20	AAK08522	NBP46 (root lectin
8	646	43.4	1434	20	AAK08530	DBX oligosaccharid
9	451.6	30.3	1981	21	AAK39492	Arabidopsis thaliana

10	441.4	29.6	1530	21	AAA66075	Potato ecto-phospho
11	164	11.0	6265	20	AAH08451	NBP6 (root lectin
12	91.2	6.1	1287	22	AAE63402	Human CD39 like pr
13	91.2	6.1	1601	21	AAE50358	Human CD39-L66 pro
14	91.2	6.1	1601	21	AAE63384	Human CDNA encodin
15	91.2	6.1	1799	21	AAE25036	Human CD39-L4 prot
16	91.2	6.1	1799	22	AAE63383	Human CDNA encodin
17	91.2	6.1	1998	21	AAAE6068	Human ecto-phospho
18	84.8	5.7	1287	21	AAE50357	Human CD39-L4 vari
19	84.8	5.7	1287	22	AAE63385	DNA encoding human
20	70.6	4.7	1574	23	ABL18793	Drosophila melanog
21	70.6	4.7	1587	23	ABL05589	Drosophila melanog
22	70.6	4.7	3901	23	ABL18792	Drosophila melanog
23	70.6	4.7	5987	23	ABL05588	Drosophila melanog
24	60	4.0	1498	22	AAE63387	cDNA encoding matu
25	60	4.0	1588	22	AAE63433	Human CD39-L2 spli
26	60	4.0	2693	22	AAE63418	Human CD39-L2 spli
27	60	4.0	2762	21	AAAE6066	Human ecto-phospho
28	60	4.0	2762	22	AAE63386	Human CD39-L2 codi
29	60	4.0	2762	22	AAE63420	Human CD39-L2 spli
30	60	4.0	2805	22	AAE63430	Human CD39-L2 spli
31	60	4.0	2882	22	AAE63424	Human CD39-L2 spli
32	56.8	3.8	2564	22	AAK54892	Human full-length
33	51.8	3.5	743	22	AAK91746	Human CDNA 5'-end
34	51.8	3.5	743	22	AAK93764	Human cDNA clone r
35	50	3.4	799	22	AAK122957	Human breast cancer
36	48	3.2	1365	21	AAAD00201	Fusion construct o
37	48	3.2	1365	21	AAAD00206	Human soluble CD39
38	48	3.2	1437	21	AAAD00202	Coding region of p
39	48	3.2	1437	21	AAAD00207	Human soluble CD39
40	48	3.2	1464	21	AAAD00204	Coding region of s
41	48	3.2	1464	21	AAAD00209	Human soluble CD39
42	48	3.2	1599	21	AAAD00200	Human soluble CD39
43	48	3.2	1599	21	AAAD00205	Human soluble CD39
44	48	3.2	1704	21	AAAE6069	Human ATP diphosph
45	48	3.2	1818	17	AAE73966	DNA coding for hum

ALIGNMENTS

ID	RESULT
AAx08529	1
AAx08529	standard; cDNA; 1489 BP

AC AAX08529

DT 19-JUL-1999 (first entry)

DE NBP46 (root lectin) cDNA.

KW NBP46; lectin; rhizobium; leguminous plant; transgenic plant.
KW nitrogen; nitrogen fixation; fertilizer; ss.

05 Lotus japonica

FH	Key	Location/Qualifiers
FM	000	05 1413

```
FT /product="NBP46 root lectin"
FT 700..702
FT misc_feature
```

FT	/note=	"these bases represent nucleotides missing
FT		from the sequence given in the specification
FT		they are included to maintain the nucleotide
FT		numbering given in the specification for this
FT		sequence"

FT /note- "These bases represent nucleotides missing
FT from the sequence given in the specification
FT They are included to maintain the nucleotide
FT numbering given in the specification for this

```

FT      misc_feature      808..810      "sequence"
FT      /*tag=d
FT      /note= "These bases represent nucleotides missing
FT      from the sequence given in the specification.
FT      They are included to maintain the nucleotide
FT      numbering given in the specification for this
FT      sequence"
XX      WO9907223-A1.
XX      18-FEB-1999.
XX      05-AUG-1998;      98WO-US16261.
XX      06-AUG-1997;      97US-0907226.
XX      (REGC ) UNIV CALIFORNIA.
XX      Etzler ME, Murphy JB;
XX      WPI; 1999-167136/14.
XX      P-PSDB; AAM85686.
XX      New polynucleotides encoding Nod factor binding lectins - useful
XX      for production of transgenic plants which are able to fix nitrogen
XX      Example 2; Pages 52-54; 57pp; English.
XX      The NBP46 root lectin is instrumental in recognising and binding
XX      to nitrogen fixing rhizobial bacteria via a lectin-carbohydrate
XX      interaction. The production of transgenic plants comprising an
XX      expression cassette expressing the NBP46 root lectin is advantageous
XX      since it would mean that non-leguminous plants could fix nitrogen
XX      from the atmosphere, lessening the need for the addition of nitrogen
XX      containing fertilizer to soil. This would lead to higher crop yields
XX      where soil has been overplanted and replenishment of the depleted
XX      soil with usable nitrogen. Alternatively, expression of NBP46 can be
XX      used to modulate oligosaccharide signalling in the plant. The nucleic
XX      acid sequences can be used to inhibit expression of an endogenous
XX      gene and also to suppress endogenous NBP46 gene expression.
XX      Sequence 1489 BP; 434 A; 284 C; 321 G; 440 T; 10 other:
SQ
Query Match      99.48; Score 1479.6; DB 20; Length 1489;
Best Local Similarity 99.38; Pred. No. 0;
Matches 1479; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
QY      1 aagtgctctctctgttagttagtgcattgagcaaaagccatgacttttaattat 60
DB      1 aagtgctctctctgttagttagtgcattgagcaaaagccatgacttttaattat 60
QY      61 ctatgaccttctgtatgtaataatgctgctatctctccccaataatctcgaaac 120
DB      61 ctatgaccttctgtatgtaataatgctgctatctctccccaataatctcgaaac 120
QY      121 aacatctaatgatagttagatattaccctcccaaaaatcagaacagttacatcatc 180
DB      121 aacatctaatgatagttagatattaccctcccaaaaatcagaacagttacatcatc 180
QY      181 gctgtatattgtagtgcagtagcaatgagaagagtcacatgctacattatcag 240
DB      181 gctgtatattgtagtgcagtagcaatgagaagagtcacatgctacattatcag 240
QY      241 aactatgatactccctccgctgaaacgaacttgattatgattcggttaaacccgt 300
DB      241 aactatgatactccctccgctgaaacgaacttgattatgattcggttaaacccgt 300
QY      301 ttgagttcaagctgctatctcgaagaagctgagaatctctgattccacttaaaa 360
DB      301 ttgagttcaagctgctatctcgaagaagctgagaatctctgattccacttaaaa 360

```

```

QY      361 gaagcagaataatgtgtctctgtgagccagcaacccaacacccgttaagcttgggca 420
DB      361 gaagcagaataatgtgtctctgtgagccagcaacccaacacccgttaagcttgggca 420
QY      421 actgcaagtttaagcgtcttggaggggaatgctgctggaataatatttgaacgctgcaag 480
DB      421 actgcaagtttaagcgtcttggaggggaatgctgctggaataatatttgaacgctgcaag 480
QY      481 gatatgctcagcaacagaagtgcccttaattgttaactcagatgacgtatctattctgat 540
DB      481 gatatgctcagcaacagaagtgcccttaattgttaactcagatgacgtatctattctgat 540
QY      541 ggaacccaagaaggttcttattcttcttgggtgacaatttaactatctcttggggaagtggga 600
DB      541 ggaacccaagaaggttcttattcttcttgggtgacaatttaactatctcttggggaagtggga 600
QY      601 aaaaatttacaagaagcagtggaagtagttagatcagagagtggtgcaatgcaatgaca 660
DB      601 aaaaatttacaagaagcagtggaagtagttagatcagagagtggtgcaatgcaatgaca 660
QY      661 tatgcagctcagaagcaacagctaaatgctccaagaatcctgaaagagagatcca 720
DB      661 tatgcagctcagaagcaacagctaaatgctccaagaatcctgaaagagagatcca 720
QY      721 tacataagaagctgttactccagggaaagaaatagacattatggttcacagttactg 780
DB      721 tacataagaagctgttactccagggaaagaaatagacattatggttcacagttactg 780
QY      781 cgtatagagaagaagcatttcgtcagagatttcaagtgctgctggttctgataat 840
DB      781 cgtatagagaagaagcatttcgtcagagatttcaagtgctgctggttctgataat 840
QY      841 ccttgattttagcgtgctttagatgggcataataatctccgggacagatataagtc 900
DB      841 ccttgattttagcgtgctttagatgggcataataatctccgggacagatataagtc 900
QY      901 tcggcccaagcttcagagatcacttgaatcaatgacaatgacaagaatagctcttaagctctt 960
DB      901 tcggcccaagcttcagagatcacttgaatcaatgacaatgacaagaatagctcttaagctctt 960
QY      961 aagatgaatgcaccttgcctccatcagaatgacatttggtagatggaatggtgga 1020
DB      961 aagatgaatgcaccttgcctccatcagaatgacatttggtagatggaatggtgga 1020
QY      1021 ggtggaagtggtcaaaaaaatcttcttacttacttacttacttacttacttacttacttactt 1080
DB      1021 ggtggaagtggtcaaaaaaatcttcttacttacttacttacttacttacttacttacttactt 1080
QY      1081 gttggatccttgtgaaataaaccatgccaatttcgtccagtgttgaagaatcgca 1140
DB      1081 gttggatccttgtgaaataaaccatgccaatttcgtccagtgttgaagaatcgca 1140
QY      1141 gctaaactagcttgaataaacaacttgaagatgcaaatcccaatcccaagatctttat 1200
DB      1141 gctaaactagcttgaataaacaacttgaagatgcaaatcccaatcccaagatctttat 1200
QY      1201 gagaagaacagtggtgaatctgtgcttgatctgtcctgaatgacacatgtgctgtt 1260
DB      1201 gagaagaacagtggtgaatctgtgcttgatctgtcctgaatgacacatgtgctgtt 1260
QY      1261 gatgatttgcttctgatacattccaagaaggttacagtggtgagaaatgaaatcag 1320
DB      1261 gatgatttgcttctgatacattccaagaaggttacagtggtgagaaatgaaatcag 1320
QY      1321 gatgctcttggaagccagcagtcgctcctgagcctgccaataagaatataatcag 1380
DB      1321 gatgctcttggaagccagcagtcgctcctgagcctgccaataagaatataatcag 1380
QY      1381 cctaaattgagagatgaatgattattttaactactagaccgtgcttaagccctggat 1440
DB      1381 cctaaattgagagatgaatgattattttaactactagaccgtgcttaagccctggat 1440
QY      1441 taccatgaagaataaataagaataaagaagccgacatcttcttcttctt 1489

```

|||||
Db 1441 taccgaagaataatcgaataaagccgacatcttcttcttcttgcctt 1489
RESULT 2
AAx08528
ID AAx08528 standard; cDNA: 1458 BP.
AC AAx08528;
XX 19-JUL-1999 (first entry)
DT 19-JUL-1999 (first entry)
XX NBP46 (root lectin) cDNA.
DE NBP46; lectin; Rhizobium; leguminous plant; transgenic plant;
XX NBP46; lectin; Rhizobium; leguminous plant; transgenic plant;
KW nitrogen; nitrogen fixation; fertilizer; ss.
XX Medicago sativa.
OS
FH Key location/Qualifiers
FT CDS complement (1..1380)
FT /*tag= a
FT /product= "NBP46 root lectin"
XX MO9907223-A1.
XX 18-FEB-1999.
XX 05-AUG-1998; 98WO-US16261.
XX 06-AUG-1997; 97US-0907226.
XX (RBCG) UNIV CALIFORNIA.
XX Etzler ME, Murphy JB;
XX WPI; 1999-167136/14.
XX P-PSDB; AAM85685.
XX
XX New polynucleotides encoding Nod factor binding lectins - useful
PT for production of transgenic plants which are able to fix nitrogen
XX
XX Example 2; Pages 49-51; 57pp; English.
PS
XX The NBP46 root lectin is instrumental in recognizing and binding
CC to nitrogen fixing rhizobial bacteria via a lectin-carbohydrate
CC interaction. The production of transgenic plants comprising an
CC expression cassette expressing the NBP46 root lectin is advantageous
CC since it would mean that non-leguminous plants could fix nitrogen
CC from the atmosphere, lessening the need for the addition of nitrogen
CC containing fertilizer to soil. This would lead to higher crop yields
CC where soil has been overplanted and replenishment of the depleted
CC soil with usable nitrogen. Alternatively, expression of NBP46 can be
CC used to modulate oligosaccharide signalling in the plant. The nucleic
CC acid sequences can be used to inhibit expression of an endogenous
CC gene and also to suppress endogenous NBP46 gene expression.
XX
SQ Sequence 1458 BP; 438 A; 268 C; 308 G; 444 T; 0 other;

|||||
Db 120 --ccaaaacagaacacctaacccttaaccccttaacgctgctatattgtagctgtagcaactgct 177
Qy 211 agcagagtcacatgcttacaattttgatcagaacttagatctctccgttgaacagaa 270
Db 178 actcgttccatcttaccatcttgcagaaacttagatcttaccatcttgcagaaact 237
Qy 271 ctgagttttatgattcgttgaacccggtttgagtttcaacgcttgaacttgaagaa 330
Db 228 attgagtttctgtaacagaatcacaacaggtttgagttgcatatgaggaataatccctgaaca 297
Qy 331 gctgcagaactctctgatacccttcaacaaagagaacgaacaaatgcttctctgagccag 390
Db 298 gacgacaaatctcttaccatcttccacttcttgaagagagaacgaagagtgcttctgaagctcg 357
Qy 391 caaccacaacacacccgttaagcttggggaacttgaacggttgaagccttttgaagggaaat 450
Db 358 caccacaacacaccccttgaagcttggggaacgaggttgaagccttttgaagggaaat 417
Qy 451 gctgctgaaaataatttgcaacgctgaggaatgctgtagcaacagaagtgcccttaact 510
Db 418 gctgctgaaaataatttgcaacgctgaggaatgctgtagcaacagaagtgcccttaact 477
Qy 511 gctcaatcagatgcagatctatctctgtaggaacccaagaaggttctatcttgggtg 570
Db 478 gctcaacgtgtagcaggttctatcttattgtaggaacccaagaaggttctatcttgggtg 537
Qy 571 acaattaaatctctcttgggaaggttgggaagaaattacaaagaacagtggaagtagt 630
Db 538 acagtttaactatgatttgggaatcttgggaagaaagcttccacaacaaatcagtggaagtagt 597
Qy 631 gatctagaaggtggttgcagtgcaaatgacatgacatgacatgacatgacatgacatgacat 690
Db 598 gaccttgaggtggttgcagtgcaaatgacatgacatgacatgacatgacatgacatgacat 657
Qy 691 gctccaaaagtaactgaaagagagatcacaatacacaataaagcttgcacccaagaag 750
Db 658 gctcctaaggtcgtgagtgagagatcacaatacacaataaagcttgcacccaagaag 717
Qy 751 aatatgaccttattgctacagttacttgcgtacatgtagaagaacattcgtgtagag 810
Db 718 caatatgaccttattgctacagttacttgcgtacatgtagaagaacattcgtgtagag 777
Qy 811 atttcaaggtcgtgagtgagtgcttgccttgccttgccttgccttgccttgccttgccttgcct 870
Db 778 gtttgaagtcacaaatgagtgcttgccttgccttgccttgccttgccttgccttgccttgcct 837
Qy 871 tatacatctcggagcagaagataagtgctcggcccgcttgcagatcactgaat 930
Db 838 ttacatctcggagcagaagataagtgcttgccttgccttgccttgccttgccttgccttgcct 897
Qy 931 caatgcagaagaatgcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 990
Db 898 gattgcagaagaatgcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 957
Qy 991 tgcacttttgggtggaatgagatgtaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1050
Db 958 tgcacttttgggtggaatgagatgtaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1017
Qy 1051 acttcaatcttattacatctccttgaaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1107
Db 1018 acttcaatcttattacatctccttgaaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1077
Qy 1108 gccaaaatctgctcagtgatttgaaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1167
Db 1078 tccatcttccatccacagtagatttgaaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1137
Qy 1168 gaggatgcagaatcccaatccacagtagatttgaaggttgaaggttgaaggttgaaggttgaaggt 1227
Db 1138 gaggatgcagaatcccaatccacagtagatttgaaggttgaaggttgaaggttgaaggttgaaggt 1194
Qy 1228 ttgagatctgctcagtgatttgaaggttgaaggttgaaggttgaaggttgaaggttgaaggttgaaggt 1287

[illegible]

Db 1163 cgagatgcgaatcctacttaccattcttgataagaanaatgtagcttcataatgta 1222
 QY 1227 ctggatcttgcttaactgtaacatctgcttgatgagatttgcttgatccattca 1286
 Db 1223 catgactcttataatcaatgtaatgctgcttgatgagatttgcttgatccattca 1282
 QY 1287 agagttacagtgagcaatgaatgaataacagatgcttgatgagagcagatggc 1346
 Db 1283 aaagattacatcaggaaggaaggaatgataccaagatgcttgatgagagcagatggc 1342
 QY 1347 tctaggaactgcatagagaacataatcatcaattgcttaattgagagatgaattat 1406
 Db 1343 tctaggaactgctgtagaagcatalcagcttaccctaattgagcagattgattat 1402
 QY 1407 tattaaact 1419
 Db 1403 tgttaagttct 1415

RESULT 5

AAF85680
 ID AAF85680 standard; DNA; 1697 BP.

AC AAF85680;

DT 25-JUN-2001 (first entry)

DE Pea blight resistance protein coding sequence #2.

KM Pea; blight resistance; nucleotide triphosphate decomposition; ds.

OS Pisum sativum.

FH Key Location/Qualifiers
 FT CDS 25..1392
 FT /tag= a
 FT /product= "Blight resistance protein"

PN JP2001017176-A.

PD 23-JAN-2001.

PF 02-JUL-1999; 99JP-0189129.

PR 02-JUL-1999; 99JP-0189129.

PA (KYOU) UNIV KYOTO.

DR WPI: 2001-320697/34.

DR P-PSDB; AAB81952.

PT New blight-resistant polypeptide useful for giving blight resistance to a plant -

PS Claim 4; Page 12-13; 20pp; Japanese.

CC The present invention provides the protein and coding sequences of a
 CC pea protein with nucleotide triphosphate decomposing activity. The gene
 CC can be used for conferring blight resistance on a plant. The present
 CC sequence is one version of the coding sequence of the invention.

XX Sequence 1697 BP; 535 A; 308 C; 306 G; 548 T; 0 other;

Query Match 56.0%; Score 833.4; DB 22; Length 1697;

Best Local Similarity 76.0%; Pred. No. 2.9e-227;
 Matches 1058; Conservative 0; Mismatches 326; Indels 9; Gaps 2;

QY 30 ttgagataagcatgactcttaattagctcatgacattggttcatgtaagc 89
 Db 12 ttacctacaacaatgaggttctcttaataactatcacttttctactatttctcatgac 71

QY 90 tgcatactctccccaatatctcgaaacaacattctcatgtaacgtaagataact 149
 Db 72 tgcataactctccccaatacttaggaacaacactcatccacagtagaagaatttct 131
 QY 150 ccccaaaaacaggaacacatcatcatcgcgtgtatattatgtagctggtgacatg 209
 Db 132 -----aaacaagaggaattctcccttaagctgtagatattcgagccgtagacatg 185
 QY 210 aagcagagtcacatgctacaattttatccaagaacttagatctctccgttgaanaag 269
 Db 186 tagtcatcatcagtttaccattttaaccagaacttagaccttctcatattgycaaag 245
 QY 270 acttggtttatgattcggttaaccggttgattgaatcagcgcataactcgttaaga 329
 Db 246 tgcagatattataaataaagataacaccgtgtttagtcatcagcctaataaccagaaca 305
 QY 330 agctcagaatctctgattccacttaaaagaagcagaanaatgltgctcctgtagcca 389
 Db 306 ggcgcagaatctctcatlccacttttagagcagaagagatgctgctccgagatct 365
 QY 390 gcaacccaacacaccggttaagcttgggcaactgcaggtttaaagctttgagagga 449
 Db 366 tcaacccaagacacccgttagacttgggcaactgcccgtttaaagcttttgaatgaga 425
 QY 450 tgcctgtaaaatataatgcaagcgttcagagatagctcagcaacaagaatgcccctta 509
 Db 426 tgccttgaagaagatattgcacatcgtaagagatagctgagcaacaagaatccttca 485
 QY 510 tgtcaatcagatgcatctatcttctgattggaacccaagaagtttctactttggt 569
 Db 486 cgttcaaccagacgcagttctcataatgattggaaccccaagaagtttctacttgggt 545
 QY 570 gacaattaactatctcttggggaagttgggaaagaattacaagaacgctggagtagt 629
 Db 546 gacagtttaactatgcatlgtggaatttaggaaagaatgacaaaacagtttggat 605
 QY 630 tgcataagagtggtgcagtgcaaatgacatgcagtcagtcagaagaacacagcttaaa 689
 Db 606 agacttggaggtgtagtcacgttcaaatggtcgtatgcagatccaagaagaacgtcaaaa 665
 QY 690 tgcctcaaaagtacctgaaaggagatccatataaagaagcttgcactcaaggaaa 749
 Db 666 tgcctcaaaagttagagatgagatgcatccatatacaaaagaagttgtaactcaaggaa 725
 QY 750 gaaatagacctttagtgcagattacttgcgcatggaagaagaagattcgagca 809
 Db 726 accataatgcttattatggttcacagttacttaccatcttggtagagaagatccgagcca 785
 QY 810 gatttcaagtcgctggtggtcgtcctaalccttgcatltagtggcttgaaggagc 869
 Db 786 gatttgaagctcacctcgttcccttaacccttgcctttagtggatttaagtgat 845
 QY 870 atatacatattccgagcagagataaaggtcgcgccccagcttaagatataactgaa 929
 Db 846 ctatacatattcagaggaagaattlaaggaacttaccacttctggtgcaaaactttaa 905
 QY 930 tcaatgcagaagatagatccttaagccttaagaatgtaacatgacttgcctatcaga 989
 Db 906 taaatgcagaacacacaaatcgtgaagccttaagttgaactatcatcatatcagaa 965
 QY 990 ttgcaacttgggtgataatggaatggtggaatggtgcaaaatcttcttcc 1049
 Db 966 ttgcaacttgggtgataatggaatggtggaatggtgcaaaatcttcttcc 1025
 QY 1050 tactcatcttcttataccctctgaagaatggtggatctt---tggatataacc 1106
 Db 1026 tcttcatcttcttcttaccacctgaagataccggtatggttgaagacacactaa 1085
 QY 1107 tgcagaatcgtccagattggttgaagaacgcgaactaaatagcttgaanaaact 1166
 Db 1086 ttctactctggcgcgcgcatatgagactaaagcttaaaagcttgcgcgttaactt 1145
 QY 1167 tgaagatgcagaatccaaatacccaagatcttattatgagaagaacagtggttgaatagtg 1226

[illegible]

RESULT 6

ID	AAF85679	standard; DNA; 1368 BP.
AA080072		

... AAF85679;

DT 25-JUN-2001 (first entry)

DE Pea blight resistance protein coding sequence #1.

KW Pea; blight resistance; nucleotide triphosphate decomposition; ds.

OS *Plum sativum.*

FH	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT /*tag=

XX

PN JP2001017176-A.

PD 23-JAN-2001.

PF 02-JUL-1999; 99JP-0189129.

PR 02-JUL-1999; 99JP-0189129.

PA (KYOU) UNIV KYOTO.

DR WPI; 2001-320697/34.

DR P-PSDB; AAB81952.

PT New blight-resistant polypeptide useful for giving blight resistance to

XX

CC The present invention provides the protein and coding sequences of a
CC pea protein with nucleotide triphosphate decomposing activity. The gene
CC can be used for conferring blight resistance on a plant. The present
CC sequence is one version of the coding sequence of the invention.

Sequence 1368 BP; 422 A; 260 C; 275 G; 411 T; 0 other;

Query Match	55.88;	Score 830.4;	DB 22;	Length 1368;
-------------	--------	--------------	--------	--------------

Matches 1049; Conservative 0; Mismatches 316;

Matches 1049; Conservative 0; Mismatches 316; Indels 9; Gaps 2;

OY 43 atgagcttccttaattagtcacatgaccttgtgtcatgtaagtgcctgcatactctcc 102
||||| ||| | |||| | | ||| | | | | ||||| ||| |||||
Db 1 atggaaattcctaattaacttaacttaacttttctaactattttctaatcctgaatcaatccc 60

QY	103	tcccaataatctcggaaacaacattctccatgaaatcgttaagataatactctcccaaaaacacg	162
Db	61	lcccaataatacttaggaaacaacactctacacgataagaagatttctct-----aaaaaca	114
QY	163	gaacacgataataacacgctgttatatttgatgcgtgtagaacgagaaagcgtccat	222
Db	115	gagaaattcttccttcacgcttagatgatttcgaatgcgcgttagacacgtgtagtccatcat	174
QY	223	gtctcaattttgtatcgaactatagatctctccctccctgttgaanaacgaaactgtttat	282
Db	175	gtttacacattttaaaccgaacttagagcctcttcattcatatttgcaaaagtgctcgatatt	234
QY	283	gattcgtgttaaaccccggttttagttcatatcagctgtcattccctgaaagacgtcgaaatct	342
Db	225	aataagataaacaacctgtgttagttcatatcagctcaataataaccagaacacgctgtcaaaact	294
QY	343	ctgattccacacttctcaaaaagagacagaaaagtggttccctgtgagccagcaaccacaaca	402
Db	295	cccatctccacattttagagcagaagagatgctgcgccgacgatcttcaaccacaagaca	354
QY	403	cccgcttaacgcttgggacaactgcagctgtttaagagcttttgaggggaaatgctgtcgtcaaat	462
Db	355	cccggttagacttgggacactgacgctccggttttaagacttttgaaaggagatgcttcgaaag	414
QY	463	atattgcaagcgtctcaaggatagctcagcaacagaagtgcctctaatgltcatatcatg	522
Db	415	atatgcatcatcgtlaaaggatagctgtagcaacagaagatccttcaacgttcaaacacagac	474
QY	523	gcacatcatctctgttattgtaacccaagaagagcttctatccttgggtggaacataacat	582
Db	475	gcaagcttctataatttgatgtaaacccaagaaggtcttctactatggtgtgacgttaactat	534
QY	583	ctctctggggaaagcttggggaaaagaattbaacaagaacagctggagtgattgataaggagt	642
Db	535	gcatctgggaaatttagggaaaaagtaacaaaacagcttgaggtaatagatcttggaggt	594
QY	643	gggtcagtgacaatgatcatatgcatgctccaaggaacacacagctaaaaatgctccaaaagta	702
Db	595	ggatcaagttcaaaatggtgcctatgcatgatacaaaaacatgctcaaaaatgctccaaaagtt	654
QY	703	cctgtaaggaagaggttccatcatataaagaactgttatcccaagggaagaataatgaaactt	762
Db	655	gaaagtgtgaagcttgcatacatcaataaagaagtgatctccaagggaatatacatatgactt	714
QY	763	tatgtctcaacttactgtgcgtatagtgaaagaaagacatttcgtgcagagatcttccaaagtc	822
Db	715	tatgtctcaacttactcaacacttggtagaagaaacatcccgagccgagattttagaagtc	774
QY	823	gctgtgtgtctgtcctaactcttgcatatttagctgtgcttgaatggtggacabatacatctcc	882
Db	775	actcctgtctctcctaacccttgccttttagcttagcgttaatgaaatcatatacatatcca	834
QY	883	ggagcagagctataaagtgctcgtgcccaacgcttcaaggtctcaacttgaatataatgcagaag	942
Db	835	gggaaagaaatttaagcgaactgcttcaactcttcgggtgcgaacacttataaatagcgaaaac	894
QY	943	atagcttctaaagcctcttaagtgaaatgcgaactgttcccatatcagaatgacacttgggt	1002
Db	895	acaattctgaagcctcttaagttgaactataccatgctcatatccagaatgacacttgggt	954
QY	1003	gggataatgaaatggttggaggtgtgaagtgtgtcaaaaaaatcttccctactcatcttcc	1062
Db	955	ggaaatttggaatggttggaggggaaatgcgacagaaaaacacttgttccttcatcttcc	1014
QY	1063	tattacccctctgaaagatggttggagctt---tgtaataaaacccaatgccaatattggt	1119
Db	1015	tttctccacacttgaagatataccggtatggttgaatgcgaacacccaattatctatactcgg	1074
QY	1120	ccaagtgtttggaagactgcgcgtcaaaactgactgtgttaaaacaacatcttgaagatgcaaa	1179
Db	1075	ccggctcgatattgagacccaagacataagaagaagcttgcggttaaaactcgagagatgcgaaa	1134
QY	1180	tccaaatacccaagatcttlatgaaagaacaggtgtgaaatattgltgtcttggatcttgc	1239

Db	1135	tctactatcatcattctctgttaagaaaaatgtaagctctcatatgatatgcatgtatcttata	1194
Oy	1240	taagttacacattgctctgtttgatgattggtgtcttgatccatctcaagaagttacagt	1299
Db	1195	tatagatgtgttactcgtttgatgattggtcttgatccatctgcaaaagattacatca	1254
Oy	1300	gcgaatgaattgatattcagaatgctgctgtgtaagacgcgaatgctccttaagcactcc	1359
Db	1255	gggaagaagaattgatattcccaagaatgctattgtgtgaagcgcgatgctccttaagcaatgc	1314
Oy	1360	atagaagcaatcatcatcattgctcctaaattgagagataatglatattattaa	1413
Db	1315	gtagaagcatatcatcagcttactctaaattgagcagttgatgtattttgttaa	1368
RESULT 7			
AA080522			
ID	AA080522	standard; cDNA; 1643 BP.	
XX			
XX	AA080522:		
DT	19-JUL-1999	(first entry)	
XX			
DE	NBP46 (root lectin) cDNA.		
XX			
KW	NBP46; Lectin; Rhizobium; leguminous plant; transgenic plant;		
XX	nitrogen; nitrogen fixation; fertilizer; ss.		
OS	Dolichos biflorus.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	51..1439	
XX		/*tag= a	
XX		/product= "NBP46 (root lectin)"	
PN	WO9907223-A1.		
PD	18-FEB-1999.		
XX			
PF	05-AUG-1998;	98WO-US16261.	
PR	06-AUG-1997;	97US-0907226.	
XX			
PA	(REGC) UNIV CALIFORNIA.		
XX			
PI	Etzler ME, Murphy JB;		
XX			
DR	WPI; 1999-167136/14.		
XX			
DR	P-PSDB; AAW85684.		
XX			
PT	New polynucleotides encoding Nod factor binding lectins - useful		
XX	for production of transgenic plants which are able to fix nitrogen		
XX			
XX	Claim 3; Page 42; 57pp; English.		
XX			
CC	The NBP46 root lectin is instrumental in recognising and binding		
CC	to nitrogen fixing rhizobial bacteria via a lectin-carbohydrate		
CC	interaction. The production of transgenic plants comprising an		
CC	expression cassette expressing the NBP46 root lectin is advantageous		
CC	since it would mean that non-leguminous plants could fix nitrogen		
CC	from the atmosphere, lessening the need for the addition of nitrogen		
CC	containing fertilizer to soil. This would lead to higher crop yields		
CC	where soil has been overplanted and replenishment of the depleted		
CC	soil with usable nitrogen. Alternatively, expression of NBP46 can be		
CC	used to modulate oligosaccharide signalling in the plant. The nucleic		
CC	acid sequences can be used to inhibit expression of an endogenous		
CC	gene and also to suppress endogenous NBP46 gene expression.		
XX			
XX	Sequence 1643 BP; 505 A; 317 C; 347 G; 474 T; 0 other;		

Best Local Similarity 76.7%; Pred. No. 9.4e-225;
Matches 1049; Conservative 0; Mismatches 309; Indels 9; Gaps 3;

Oy	51	cttaataagctcctaigaaccttctgttcaatgtaaaagcctgctactcttctcccaata	110
Db	86	cttctctactccctacccctcttcttactctcttccatctgcaaaaacttctcttctgcgaata	145
Oy	111	tctcgcgaagaacacttctcgtatgctgtaagaatbatactaccaccaaaaatacgaagaacct	170
Db	146	tgttgcggagaacagatcttactaatactcgttaagaatbactcccaac---caggaaacctct	202
Oy	171	taactcatagcgcgttatatttgatctgtctgttggaacttgaaacagagccactgtctcaaa	230
Db	203	tacctcttaagcgtctactcttctgtatgctgtgctctcctggagctgctgcacatgcttcaaa	262
Oy	231	ttttgcacagaactatgatctccctccgtttgaaaaacgaacttgattatgctgt	290
Db	263	ttttgcacagaaacttatgactctctcgtacattgacatgacatgacctgcgtgtttacaaaaagat	322
Oy	291	taaacccggttttgatgttcaatacgcctgtaactcctgaaagaagcttgaaagatctctgattcc	350
Db	323	caaacccggttttgatgcctacacgtctgataagcttgaaaaagcttgaaatctctccatct	382
Oy	351	actctctaaagaagcagaaaatgttgctctgtgaaocagacaccaacacaccggttaa	410
Db	383	acttttgagagaaagctggaagatgtgtccctctgaaagaaactgcaccccaagaaccccttaa	442
Oy	411	gctttggggcaactgcaggtttaaagctcttttgaaagggttaagctgtctgtaaaaataattgca	470
Db	443	gctttggggcaacacagcaggttttgagtgctcttgtaatgggtgtctgtctgtaaaaagatttgc	502
Oy	471	agcgcataaggaatatgctcgcacacaggaagctgcccataagcttgaatacagaatgcagatc	530
Db	503	agcgcgttaagggaatgttccagagacacgaagaagcttccctgagcgttcaactgatatgcagatc	562
Oy	531	tattcttgatgtaaaccccaagaaggtctcttaactcttgggtgtaaccaatlaactatccctggg	590
Db	563	tgttatttgatgtaaaccccaagaaggtctcttaactatgattgttcaagtttaactatctgtttaa	622
Oy	591	gaagtctggaaaaaagatttcaaaagcagtggaagtagtattgatacagaagctgtgtcagat	650
Db	623	aaagtctggaaaaaagatttcaaaactgtgggaagtgtataagctcttggaaagtgtctcagat	682
Oy	651	gcaaatgacatactgcaagctcgaaggaacaacagctaaaaaatgctccaaaaagtactctgaag	710
Db	683	tcaaatgcttatctgtctcgaagaataacagctaaaaaatgctccaaaaacaaccaacaag	742
Oy	711	agaagatccctaataaaagagcttcttaacgaagggaagaataatgtaacttataatgttca	770
Db	743	agaagatccctaataaaagagcttcttaacgaagggaagaataatgtaacttataatgttca	802
Oy	771	cagttacttgcgtatgtaagaagaagcattctgtgcaagaaatttcaagatgcgtcgtg	830
Db	803	cagttacttgcgtatgtaagaagaagcgtgttaagatttttaagaccactatg	862
Oy	831	ttctgcataactcttgcattttaagctgtgctttagatggggaatataacttaatttcoggaacga	890
Db	863	tgtcgtcagaactcctgtctatttgcagcgtatgaagaatatatacagatatctccgagaatc	922
Oy	891	gtataaagctctgcgccccacagctcagaactcaatgtaacatgcagaagaatagcct	950
Db	923	gtataaatactatctatgttccactctctgtgccaacttaatggtgcgtgcgacctagctct	982
Oy	951	taagagctcttaaaagtgaatgcaccttctccctactcagaatgcactgtcactttgtgtggagatg	1010
Db	983	tcaagatctcagaattgaatgaatgacatgttccatgaaaaactgcacacttgtgtggagatg	1042
Oy	1011	gaatgctggaggggtggaagctgtcaaaaaaactcttctcacttactatcttcttactact	1070
Db	1043	ggaatgctggagaaaggaagtgtgacagaaaaaactctgtgttactctcagcttctcactacta	1102
Oy	1071	ctctgaagaatgttggatcttctgtgataaacaacatgcaaaattgcgtlccagttgattt	1130

Db 1005 tggcgtgctggtgagctggggaacaacttttctgtcatcatttcttgaagtcgc 1064
Qy 1074 tgaagatgttgagctcttctt---tgaataacccaataattgctcagttgttc 1130
Db 1065 cgaatgagccggtgttctgtcatccaacgacgcatgctcctctgtgatt 1124
Qy 1131 gaagactgcagctaaactagctgttgaacaactctgagatgcacaataccc 1190
Db 1125 tgaagatgcagcaaggttctgtctgtacacagaattaaagatctcaagcgtttccc 1184
Qy 1191 agactcttagagaagaacacgctgtgatatagtgtctgtgattcttctaagttaac 1250
Db 1185 tcggttaaggaatgaga---tcttcctacatagttctgtgattagatataccaattac 1241
Qy 1251 attgctgttgatgattgtctgtcatcattccaagaagttacaagttgcagatgaat 1310
Db 1242 attgctgttgatgattgtctgtcatcattccccaagaagattacattggtgagccaat 1301
Qy 1311 tgaatatacagatgctcttctgtgaagccgcatgtgcctctagcactgacataagaacat 1370
Db 1302 tcaatatacagatctctctgtgagctgcatgcccactgcgaagttgcataagaagcat 1361
Qy 1371 atcatcttgccctaaatttgagagataatgtatttatttaact 1416
Db 1362 atctctgtacctaatttgagaataattgtatttcttctaagct 1407

RESULT 9

AAC39492
ID AAC39492 standard; DNA; 1981 BP.

AC AAC39492;

DT 17-OCT-2000 (First entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 24831.

KW Hybridisation assay; genetic mapping; gene expression control;

KM protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS EP1033405-r2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 11-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142154.

PR 06-JUL-1999; 99US-0142055.

PR 08-JUL-1999; 99US-0142390.

PR 09-JUL-1999; 99US-0142803.

PR 12-JUL-1999; 99US-0142920.

PR 13-JUL-1999; 99US-0142977.

PR 14-JUL-1999; 99US-0143542.

PR 15-JUL-1999; 99US-0143624.

PR 16-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 19-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 20-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

Db 3239 aggt 3242

RESULT 12

ID AAF63402 standard; DNA; 1287 BP.

AA63402;

14-MAY-2001 (first entry)

Human CD39 like protein CD39-L4 coding sequence.

Human CD39-like protein; apyrase; NDase; platelet function inhibitor; myocardial infarction; cerebral ischaemia; angina; arterial thrombosis; cerebral artery thrombosis; platelet aggregation; inflammation; apoptosis; autoimmune disorder; neurological disorder; Alzheimer's disease; Parkinson's disease; cancer; CD39-L4; ds.

Homo sapiens.

WO200110205-A1.

15-FEB-2001.

09-AUG-2000; 2000WO-US21790.

09-AUG-1999; 99US-0370265.

11-JAN-2000; 2000US-0481238.

25-APR-2000; 2000US-0557800.

26-MAY-2000; 2000US-0583231.

30-JUN-2000; 2000US-0608285.

(HYSE-) HYSEQ INC.

Ford J, Mulero JJ, Yeung G;

WPI; 2001-147489/15.

P-PSDB; AAB72243.

Polynucleotides encoding human CD39-like polypeptides, with apyrase and/or NDase activity, which are useful in the treatment of pathological conditions caused by thrombosis (e.g. myocardial infarction) and inflammatory disorders -

Disclosure: Page 140-142; 203pp; English.

This invention relates to polynucleotides encoding human CD39-L4-like polypeptides with apyrase and/or NDase activity. The polypeptides having apyrase, including NDase, activity are useful for inhibiting platelet function and can therefore be used in the prophylaxis or treatment of pathological conditions caused by or involving thrombosis or excessive coagulation or excessive platelet aggregation, such as myocardial infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral artery thrombosis or intracardiac thrombosis, and conditions associated with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in modulating disease states (including platelet aggregation, inflammation and apoptosis) associated with ADP or other purinergic signalling by reducing the levels of NDPS. The polypeptides are also useful for prophylaxis or treatment of inflammation related disorders, such as disorders involving sepsis or systemic inflammatory response syndrome or SIRS (and associated conditions such as fever, tachycardia, tachypnea, cytokine overstimulation); autoimmune disorders such as thrombosis, atherosclerosis, acute pancreatitis, dermatitis, including psoriasis, cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis; neurological disorders including neurodegenerative diseases, epilepsy, depression, Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis; and cancer. The present sequence represents the CD39 like protein CD39-L4 coding sequence.

Sequence 1287 BP; 330 A; 290 C; 344 G; 323 T; 0 other;

Query Match 6.1%; Score 91.2; DB 22; Length 1287; Best Local Similarity 50.0%; Pred. No. 14e-15; Matches 256; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY	150	ccccaaatcaggaacccggtacatcacgcgtgtatatttgatgctgtagcaactg	209
DB	117	ccccatcattgtagcgcgcacacgtgtatgattgattgtagcaggagcaactg	176
QY	210	aagcagatccatgctcacaatttgatacgaactagatctctccgttgaaacga	269
DB	177	aactcgattcatgtttacaaccttgtagaanaatgcagacagcttccaaattaga	236
QY	270	acttgagtttatgattgctggttaaccggtttgagttcatacgtcgtcaatctga	329
DB	237	agggaagttttgattgctgtagaagcagacacttctgcttttgatgaactaagca	296
QY	330	agctgcagaaatctctgattccacttctaaagaagaagaatggtgcttgtagcca	389
DB	297	gggtgctgagaccgttcaagggctcttagaggtggtccaaagactcaatcccgaa	356
QY	390	gcaaccacaacacccgttaagcttggtggaactgcaggtttaagctttggaggaa	449
DB	357	ctggaanaaagaccgccagtggtcttaaggaacagagagactacgttactgcaga	416
QY	450	tgctgtgaaatataattgcaagcggtcagggatgctcagcaacagaagtgccctla	509
DB	417	caagccaaaggtctctgctcttgagttaaagagatcttcaggaaagcttccctc	473
QY	510	tgctcaatcagatgcagatctatcttcttgatggaaccgaagaagttcttattg	569
DB	474	ggtcaaaagggaggttagatgacatgagatggtacgcgaagaagcatatagctg	533
QY	570	gacaattactatctctcttggtggaagtgggaattcaacagaacagtggaagtag	629
DB	534	tactgtgaattcttcgacagatgcagctgcatgccaagaacagagagactgtgg	593
QY	630	tgatctagaggttgggtcagtgcaaatgacat	661
DB	594	ggactagggggagccctccacccaatcacgt	625

RESULT 13

AAZ50358

ID AAZ50358 standard; CDNA; 1601 BP.

AAZ50358;

18-MAY-2000 (first entry)

Human CD39-L66 protein encoding CDNA.

CD39-L4; human; CD39-L66; apyrase; nucleotide diphosphatase; NDase; ADP diphosphohydrolase; Apyrase; adenosine diphosphate; ADP; treatment; platelet aggregation; antithrombotic; thrombosis; myocardial infarction; cerebral ischaemia; angina; vascular graft; extracorporeal circulation; molecular weight marker; nutritional supplement; tumour; prevention; drug targeting; splice variant; ss.

Homo sapiens.

Location/Qualifiers

Key 246..1463

CDS /*tag= a

/product= "Human CD39-L66 protein"

/note= "Splice variant of CD39-L4 protein"

WO200004041-A2.

27-JAN-2000.

16-JUL-1999; 99WO-US16180.

PR 16-JUL-1998; 98US-0118205.
 PR 24-JUL-1998; 98US-0122449.
 PR 04-FEB-1999; 99US-0244444.
 PR 19-MAR-1999; 99US-0273447.
 PR 09-JUL-1999; 99US-0350836.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford J, Mulero J;
 XX
 DR WPI: 2000-182397/16.
 DR P-PSDB: AAF44851.
 XX
 PR New nucleic acid encoding human CD39-like protein, useful for treating
 PI and preventing thrombotic disease -
 XX
 PS Claim 8; Page 123; 125pp; English.
 XX
 CC The present sequence is a cDNA encoding the CD39-like protein, a splice
 CC variant of CD39-L4 protein. It is an apyrase and/or nucleotide
 CC diphosphatase (NDPase), isolated from the human foetal liver-spleen cDNA
 CC library, b2HFLS20W. It is a soluble ATP diphosphohydrolases (ATPase) and
 CC is involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
 CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%
 CC homology to human and murine CD39. It has platelet aggregation inhibition
 CC and antithrombotic activity. CD39-L4 is used to treat or prevent
 CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
 CC also used in vitro, to maintain vascular grafts or during extracorporeal
 CC circulation, to hydrolyse NDP, as molecular weight markers and as
 CC nutritional supplements. It is used to identify therapeutic agents that
 CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs
 CC to tumours or other cells that express CD39-L4.
 XX
 SO Sequence 1601 BP; 436 A; 352 C; 428 G; 385 T; 0 other;

Query Match 6.1%; Score 91.2; DB 21; Length 1601;
 Best Local Similarity 50.0%; Pred. No. 1.5e-15;
 Matches 256; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY 150 ccccaaaatcaggaaccagttacatcagctgttatatttgcgtgtagcactg 209
 DB 362 ccccatcagtcagcagcagcactgtatgtaattgttgcagggagcactg 421
 QY 210 aagcagagtcacgtctcaatttgcagaacttagctccctccgttgaaacga 269
 DB 422 aactcgaattcagttacaccttgcagaaatgcagacagcttccaattctaga 481
 QY 270 acttgatttattgattcgttaaacccggttgcattacatcagctcctaga 329
 DB 482 aggggaagtttgcattcgtgaagccagacttctgttgcatacactaaga 541
 QY 330 agctcagagatcctgattccaactttaaagaagcagaanaatggtctcgttgaca 389
 DB 542 gggctcgtgacgcgttcaagggtcttagagtgccaaagctcaatcccccagatca 601
 QY 390 gcaacccaacacccggttaagcttgggcaactgcaggtttaaagcttggagggaa 449
 DB 602 ctggaanaagagccagtgcttcaagagcagcagacacttctcccaaga 661
 QY 450 tgcctcgtgaataatattcgaagcgttcaggagatgctcgaacacagagtgccctta 509
 DB 662 caaagcgaagctcgtcttgcagtaagaagatcttcagaagaagctccttctt--- 718
 QY 510 tcttaacatcagatgagatctatcttctgaagcaacccaaggttcttattcttgggt 569
 DB 719 ggtacccaagggcagtgtagcatgcatgcatcgcagcagagatattgacttgggt 778
 QY 570 gacaaatcattctcttcttggggaagttggaaaagattacaaagacagtggtgagta 629
 DB 779 tacttggaatttctctgacagtcagctgcatggtccacagacagagacgttggtggagctt 838
 QY 630 tgactagagagtggtgctcagtgcaaatgacat 661

DB 839 ggacttagggggagcctccaccacaatcagct 870
 || ||||| || | || ||||| || |
 RESULT 14
 ID AAF63384
 AC AAF63384 standard; CDNA; 1601 BP.
 XX
 AC AAF63384;
 XX
 DT 14-MAY-2001 (first entry)
 XX
 DE Human cDNA encoding CD39 like protein CD39-L4.
 XX
 KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
 KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
 KW cerebral artery thrombosis; platelet aggregation; inflammation;
 KW apoptosis; autoimmune disorder; neurological disorder;
 KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200110205-A1.
 XX
 PD 15-FEB-2001.
 XX
 PF 09-AUG-2000; 2000WO-US21790.
 XX
 PR 09-AUG-1999; 99US-0370265.
 PR 11-JAN-2000; 2000US-0481238.
 PR 25-APR-2000; 2000US-0557800.
 PR 26-MAY-2000; 2000US-0583231.
 PR 30-JUN-2000; 2000US-0608285.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford J, Mulero J, Yeung G;
 XX
 DR WPI: 2001-147489/15.
 DR P-PSDB: AAF72239.
 XX
 PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
 PT and/or NDPase activity, which are useful in the treatment of
 PT pathological conditions caused by thrombosis (e.g. myocardial
 PT infarction) and inflammatory disorders -
 XX
 PS Claim 10; Page 156-157; 203pp; English.
 XX
 CC This invention relates to polynucleotides encoding human CD39-like
 CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
 CC Arnpase, including NDPase, activity are useful for inhibiting platelet
 CC function and can therefore be used in the prophylaxis or treatment of
 CC pathological conditions caused by or involving thrombosis or excessive
 CC coagulation or excessive platelet aggregation, such as myocardial
 CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
 CC artery thrombosis or intracardiac thrombosis, and conditions associated
 CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
 CC modulating disease states (including platelet aggregation, inflammation
 CC and apoptosis) associated with ADP or other purinergic signalling by
 CC reducing the levels of NDPs. The polypeptides are also useful for
 CC prophylaxis or treatment of inflammation related disorders, such as
 CC disorders involving sepsis or systemic inflammatory response syndrome or
 CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
 CC cytokine overstimulation); autoimmune disorders such as thrombosis,
 CC atherosclerosis, acute pancreatitis, dermatitis. Including psoriasis,
 CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
 CC neurological disorders including neurodegenerative diseases, epilepsy,
 CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
 CC disease, and amyotrophic lateral sclerosis; and cancer. The present
 CC sequence represents cDNA encoding CD39 like protein CD39-L4.
 XX
 SO Sequence 1601 BP; 436 A; 352 C; 428 G; 385 T; 0 other;

THIS PAGE BLANK (USPTO)

OY	150	cccccaaatcaggaacacagttacatactacagtcgttatattgatgtgtgaacacgg	205
Db	318	ccccattaatctcagtcgcccgcacaccttcttaagaaattgtttgatctggggcaacacgg	373
OY	210	aagcagagtcacatgctctacaattttgatacagaacttaagatctctctcccgcttgaanaacga	265
Db	378	agctcggattcatgttttactactctttgtgcagaaaacagcagagaacgcccctctcttgcga	433
OY	270	acttgagtttatactgattccgttttaaacccggttttgatgtcatacgtctctaaactctgaaga	325
Db	438	aggttgaaatttttctgtctcttgaaagcccgcgagacttctcttcttgatctagccacaaca	493
OY	330	agcttcgaaatctctgattccactctcttaagaagaacagaaatggtgtctcttgtagggcca	385
Db	498	gggtgtctgagactgtccagagagctctcttvgaggtgcccacaagactcgatcccacagaagcca	553
OY	390	gcaaacccacacaccccgcttgaagctcttvgggcagacactgcagtttaaggcttttggaggggaa	445
Db	558	ctgggaagaagaccccggtgtgtcttgaagaagcagccgagactcgtttgtctgcttgagca	613
OY	450	tgctgctgtaaaatatattgcgaagcgggtcagagatatgtctcagcaacagaaatgcctctaa	505

Db 618 gaagccagctctgctcttgagtagagagatctcaagaattcccttctc--- 674
QY 510 ttttaataatagatcagatctatcttcttgatgaaacccaagaagttcttactcttggt 569
Db 675 ggtcccaagatgacagcttagatcatgatgaggtcctctatgaaagcatalactagcctgggt 734
QY 570 gacattactctctcttgagggaattgggaaagatttacaagacagctggagtggt 629
Db 735 taacgtgaactcttcttaacaaagtcagctgcatgctgctggccagagagactgtgggaccct 794
QY 630 tgatctagaggtgggtcagtgcaatgacatat 663
Db 795 tgacctgggggtgcctccacccaataacagcttt 828

RESULT 2
US-09-608-285A-4
Sequence 4, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608, 285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583, 231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557, 800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481, 238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370, 265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350, 836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273, 447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244, 444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122, 449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118, 205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 4
LENGTH: 1287
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1284)
US-09-608-285A-4

Query Match 6.1%; Score 91.2; DB 4; Length 1287;
Best Local Similarity 50.0%; Pred. No.1.5e-18;
Matches 256; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY 150 ccccaaaaatcaggaacagttacatcatcagctgtatataattgagtctgtagacatg 209
Db 117 ccccatcaatgtaacgagcagcactgtatgtgaattatgttgatgacagggagcactg 176
QY 210 aagcagagtcacatgcttaacaatttgatcagaacttagatctctcccgltgaaacga 269
Db 177 aactcgaattacatgcttaacaccttgtgcagaaatgccagcagcttccaacttaga 236
QY 270 acttgagttatgatctcggttaaacccggttgagttcattacgcgtcataactctgaaga 329

Db 237 aggggaagtttttgattctgtgaagccagagacttcgtctttgtagatcaactaagca 296
QY 330 agctgcgaactctctgtattccacttcttaaaagaacgaagaattgtgtctctgagcca 389
Db 297 ggggtgtgaagccgtltaaaaggtccttagagtggtgccaagaactcaatccccgagtca 356
QY 390 gcaaccacacacaccgcttaagcttgaggcaactgcaggtttaagcttttgaggga 449
Db 357 ctggaanaagacccccagttgtcttaaaaggcacaacgcagcagactaactgccaagaca 416
QY 450 tgcgtctgaataatatttgcaagcgtcagggatagatctagaacaagaatgcccctaa 509
Db 417 caagccaagctctgctcttgagtgaaagagatctcagaagatccacacttctc--- 473
QY 510 tgtcaatcagatcagatcatctatcttgatgaaacccaagaagttcttactcttggt 569
Db 474 gttaccaagggcagctgtgacatcagatgatalgataccgacgaagacatatlaacttggt 533
QY 570 gacattactatctcttgagggaagttgggaaagatttacaagaacagctgggagtagt 629
Db 534 tactgtgaatttctgacaggtcagctgcagatgcccacagacagagagactgtgggacctt 593
QY 630 tgatctagaggtgtgctcagtgcaatgacat 661
Db 594 ggaactaggggagcctccacccaataacagct 625

RESULT 3
US-09-608-285A-24
Sequence 24, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608, 285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583, 231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557, 800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481, 238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370, 265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350, 836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273, 447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244, 444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122, 449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118, 205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 24
LENGTH: 1601
TYPE: DNA
ORGANISM: Homo sapiens
US-09-608-285A-24

Query Match 6.1%; Score 91.2; DB 4; Length 1601;
Best Local Similarity 50.0%; Pred. No.1.7e-18;
Matches 256; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

```

1 CURRENT APPLICATION NUMBER: 2000-06-30
2 CURRENT FILING DATE: 2000-05/09/608,2685A
3 PRIOR APPLICATION NUMBER: 09/583,231
4 PRIOR FILING DATE: 2000-05-26
5 PRIOR APPLICATION NUMBER: 09/557,800
6 PRIOR FILING DATE: 2000-04-25
7 PRIOR APPLICATION NUMBER: 09/481,238
8 PRIOR FILING DATE: 2000-01-11
9 PRIOR APPLICATION NUMBER: 09/370,265
10 PRIOR FILING DATE: 1999-08-09
11 PRIOR APPLICATION NUMBER: PCT/US99/16180
12 PRIOR FILING DATE: 1999-07-16
13 PRIOR APPLICATION NUMBER: 09/350,836
14 PRIOR FILING DATE: 1999-07-09
15 PRIOR APPLICATION NUMBER: 09/273,447
16 PRIOR FILING DATE: 1999-03-19
17 PRIOR APPLICATION NUMBER: 09/244,444
18 PRIOR FILING DATE: 1999-02-04
19 PRIOR APPLICATION NUMBER: 09/122,449
20 PRIOR FILING DATE: 1998-07-24
21 PRIOR APPLICATION NUMBER: 09/118,205
22 PRIOR FILING DATE: 1998-07-16
23 NUMBER OF SEQ ID NOS: 60
24 SOFTWARE: Patentin Ver. 2.0

```

OY	150	ccccaaaatcagaagcaacgattacatacatagctgttatattttgagtgctgtgacactg	209
Db	362	cccccaaatgtcagcgccgacgacctgttatgatataattgttatgttcagggagcacttg	421
OY	210	aagcagagtcacatctacataattctgtgtcgaacttgatctccctccgttgaacaag	269
Db	422	aaccggaattcatgttttaacaccttcttgccaggaatatgcagagacgttccaattctaga	481
OY	270	acttgagtttatagtatccgtttaaaccggttggagtcatcacgctgtcaatccctgaag	329
Db	482	agggagaagtttttattctctggaagccagagactcttcgctcttgttaatatcaactaaga	541
OY	330	agcgcagaaatctctgattcccaactcccaagaagacagaaatgtggtctccttgagcca	389
Db	542	gggtgctcgtgagccgttcaagggctcttaagggtggtgccaaagactcaatcccccgaagta	601
OY	390	gcaaccaccaacaccccggttaagctcttggtgccaactgcaggttttaagctttttagaggaa	449
Db	602	ctggaaaaaagcccccagttgtctctaaaggcaacgcagagactacgcttaattgcagaaca	661
OY	450	tgtcgtctgaanaatatctgcacaagcgctcagaggaatatctcagacaacgaagtgcccttaa	509
Db	662	caaaagccaaagctctgcctcttcttggtgflaaaggagatctctcagaagagtcaccttccc	718
OY	510	tgttcaatcagatgcagtatctattcttgatggaaacccaagaaggtcttatctattcttggt	569
Db	719	ggtcccaaaaggccagtglttaagcatcaatcagatgagatcccgacgaagcatataaagcttgg	778
OY	570	gacaaatbaacatctctctgggaagttctgggaaaaagaattacaagaacagtggaagagt	629
Db	779	tacgtgtgaattctctgcacaggtccagctgcgatgcacaaagacagagagactgttggagacct	838
OY	630	tgatctaggaaggtggtctcaatgcgaatgacat	661
Db	839	ggacctataggggagctccacaaacatcacagt	870

```

RESULT 5
US-09-240-639-5
; Sequence 5, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischauf, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; FILE REFERENCE: POLYPEPTIDES AND NUCLEIC ACIDS
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ. ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```



```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-608-285A-45
```

```

Query Match      4.0%; Score 60; DB 4; Length 1498;
Best Local Similarity 48.1%; Pred. No. 9,8e-09;
Matches 233; Conservative 0; Mismatches 245; Indels 6; Gaps 2;
```

```

OY 178 taagcgtatataattatctgtgtagcactgtgaagcagatccatctctacaatttgat 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 304 taaggatcatatgtttgtagcagaacacgtgcagccggtacacgtctccagttcacc 363
OY 238 cagaacttagatctccctccgttgaacacgaacttgatttatatttcogttaaacc 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 364 cggccccccagagaacactccacgttaacc--cacgaacacttcaagcagtgaaagcca 420
OY 298 ggtttagttcatagcctctaaactctgaagaagctgcagaatctctgattccattcta 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 ggtcttctgcctatgtcatgtatgttgaagaagcgtcgaaggaatccgggaactctg 480
OY 358 aaagaagcagaagaatgtgttctctgtgagccagcaaccacacccgttaagcttggg 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 gatgtgtctaaacagacacttccgttcgacttctgtgaagggccacccctctggtctcaag 540
OY 418 gcaactgcaggtttaaggcttttgaggggaatgctgctgaataatatattgcaagcgttc 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 gccacagcgtgcttaagcctgttaactctggaagaagccacagaagttaactgcagaaggtg 600
OY 478 agggatattgtcagcaacagaagtgcccttaatgttcaatcagatcagatattctt 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 601 aaagaagattttaagcattgccttccctctgttaggggat--gactgtgttccatcatg 657
OY 538 gatggaaaccagaagaagttcttattcttgggtgacaattacattctcttggggaagttg 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 658 aagggagacagatgaagcgcttgcggcgtgatccatcacttccctgaagcagcaagttg 717
OY 598 ggaanaagatttacaagaagcgtgggagtagttgatctagaaggtgtgtcagcgaagtg 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 718 aaacaccacaggaagcagcgtggcagctgctggaacttggcggaagatccactcagatc 777
OY 658 acat 661
    ||
DB 778 gcct 781
```

```

RESULT      8
US-09-608-285A-53
; Sequence 53, Application US/09608285A
; Patent No. 6335013
```

```

; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-608-285A-53
```

```

Query Match      4.0%; Score 60; DB 4; Length 1588;
Best Local Similarity 48.1%; Pred. No. 1e-08;
Matches 233; Conservative 0; Mismatches 245; Indels 6; Gaps 2;
```

```

OY 178 taagcgtatataattatctgtgtagcactgtgaagcagatccatctctacaatttgat 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 taaggatcatatgtttgtagcagaacacgtgcagccggtacacgtctccagttcacc 510
OY 238 cagaacttagatctccctccgttgaacacgaacttgatttatatttcogttaaacc 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 511 cggccccccagagaacactccacgttaacc--cacgaacacttcaagcagtgaaagcca 567
OY 298 ggtttagttcatagcctctaaactctgaagaagctgcagaatctctgattccattcta 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 568 ggtcttctgcctatgtcatgtatgttgaagaagcgtcgaaggaatccgggaactctg 627
OY 358 aaagaagcagaagaatgtgttctctgtgagccagcaaccacacacccgttaagcttggg 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 628 gatgtgtctaaacagacacttccgttcgacttctgtgaagggccacccctctggtcccaag 667
OY 418 gcaactgcaggtttaaggcttttgaggggaatgctgctgaataatatattgcaagcgttc 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 688 gccacagcgtgcttaagcctgttaactctggaagaagccacagaagttaactgcagaagtg 747
OY 478 agggatattgtcagcaacagaagtgcccttaatgttcaatcagatcagatattctt 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 748 aaagaagattttaagcattgccttccctctgttaggggat--gactgtgttccatcatg 804
OY 538 gatggaaaccagaagaagttcttattcttgggtgacaattacattctcttggggaagttg 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 805 aacggaaacagatgaagcgcttgcggcgtgatccatcacttccctgaagcagcagcttg 864
OY 598 ggaanaagatttacaagaagcgtgggagtagttgatctagaaggtgtgtcagcgaagtg 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 865 aaacaccacaggaagcagcgtggcagctgctggaacttggcggaagatccactcagatc 924
```

OY	178	tacgcigtatatttgatgctcgtgtagaacctggagacgaatccatgctacatatttgat	237
Db	382	tacggagatcatgctttgttgtagcaggaagaaacctggtcccccgaataacagcttcccgattacc	441
OY	238	cagaactatgatctctccctccggttgaaacacgaacttgatttatgattcgtgttaacc	297
Db	442	cgtccccccagagaactccacgaacttacc---cagaaaccttcaagaagctgaagca	498
OY	298	ggttttgatcatatagcgtctgcttaaccctggaagaagctctgaaatctcgtatccattcta	357
Db	499	ggtcttctcgtcctatgctcatgatgatgttgaaagaagcgtctgaaacacccgggaactctg	558
OY	358	aaagaagcagaanaatgtgcttctgtgtagccagcaaccacacaccccgcttaagcttg	417
Db	559	gatttcttaacacagagacattccgtctcgaattctlgaaagccacccctctgctccoaag	618
OY	418	gcaactcagatttaagaagcttttgaggaggaaatgctctgtaaaatatattgcaagaagctc	477
Db	619	gccacagctggtcttaccgcttaccctcgtgagcgggaaaaagcccaagaattactctgcgaagctg	678
OY	478	aggagatagctcagcaacagaaatgycctctaattgatttcaatcagaatgagtatattctt	537

511

238 cagaacttagatctcctccgttgaaacgaactgagtttatgatcggttaacc 297
| | | |||| | | | | | | | |
511 cgagcccccaagaacactcccacqftaac---cacgaacctcaaagcaqftaaqcga 567

Db 511 cggccccagagaactccacgyltaac---cacgaaccttcaaaagcagtgaaagcca 567
Qy 298 ggttgatcatacgtctgaatccctgaagaagctgcagatatctctattccacttcta 357
Db 568 ggtcttctgctatgctgcatgcatgtttgaaagagcgtctcaggaatccggaactactg 627
Qy 338 aaagaagcagaanaatgtgttctctctgtgagcagcaaccaacacacccgttaagcttgg 417
Db 628 gatgtgttaacagagacatccgtctgacttctggaagcgacacccctctgtctctcaag 687
Qy 418 gcaactgcaggttaagcttcttggaaggaatgctgctgtaaatatatttcaagcgctc 477
Db 668 gccacagctgcttaagcctctgttaactgtgaaagagcccaaggttaactgcagaagtg 747
Qy 478 agggatgctcagcaaaagagatgccccttaattgtcaatcagatgcagttacttct 537
Db 748 aaagaagatttaagacatcgcttctctctgttagggat---gactgtgtttccatcatg 804
Qy 538 gatggaaccagaaaggtcttctatcttgggtgacaaatattctcttggggagtg 597
Db 805 aacggaacagatgaagcggttctgcggtggaaccacatcaacttctctgacagcgactg 864
Qy 598 ggaagaagattacaagaagcagtggaagtgatgactagtaggtgtggtcagtgcaaatg 657
Db 865 aaactccagagagggagcagcggtggcatgctgacttggcgaggaatccactcagatc 924
Qy 658 acat 661
Db 925 gcct 928

RESULT 13
US-09-608-285A-50
; Sequence 50, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO: 50
; LENGTH: 2805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-608-285A-50

Query Match 4.0%; Score 60; DB 4; Length 2805;
Best Local Similarity 48.1%; Pred. No. 1.4e-08;
Matches 233; Conservative 0; Mismatches 245; Indels 6; Gaps 2;
Qy 178 taagctgttatattgactgtgtaagcactggaagcagatgcttactaaatttgat 237
Db 451 taaggatcatgtttgtgacgaagacactgcacccagatatacagcttccagttcaac 510
Qy 238 cagacttaatctccctcccggttgaagaagacttgagttatgatttggttaaac 237
Db 511 cggccccagagaactccacgyltaac---cacgaaccttcaaaagcagtgaaagcca 567
Qy 298 ggttgatcatacgtctgaatccctgaagaagctgcagatatctctattccacttcta 357
Db 568 ggtcttctgctatgctgcatgcatgtttgaaagagcgtctcaggaatccggaactactg 627
Qy 358 aaagaagcagaanaatgtgttctctctgtgagcagcaaccaacacacccgttaagcttgg 417
Db 628 gatgtgttaacagagacatccgtctgacttctggaagcgacacccctctgtctctcaag 687
Qy 418 gcaactgcaggttaagcttcttggaaggaatgctgctgtaaatatatttcaagcgctc 477
Db 668 gccacagctgcttaagcctctgttaactgtgaaagagcccaaggttaactgcagaagtg 747
Qy 478 agggatgctcagcaaaagagatgccccttaattgtcaatcagatgcagttacttct 537
Db 748 aaagaagatttaagacatcgcttctctctgttagggat---gactgtgtttccatcatg 804
Qy 538 gatggaaccagaaaggtcttctatcttgggtgacaaatattctcttggggagtg 597
Db 805 aacggaacagatgaagcggttctgcggtggaaccacatcaacttctctgacagcgactg 864
Qy 598 ggaagaagattacaagaagcagtggaagtgatgactagtaggtgtggtcagtgcaaatg 657
Db 865 aaactccagagagggagcagcggtggcatgctgacttggcgaggaatccactcagatc 924
Qy 658 acat 661
Db 925 gcct 928

RESULT 14
US-09-608-285A-54
; Sequence 54, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

query Match      3.38; Score 49.8; DB 1; Length 7218;
Best Local Similarity 5.88; Pred. No. 4e-05;
Matches 24; Conservative 216; Mismatches 173; Indels 0; Gaps 0

```

Oy	444	ggggaatgcctcgaaataatttgcgaacggtcgaaggtatctgcacgaacaagaatgc	503
Db	1434	RRR	13757
Oy	504	cctaattgtccaatcagatgcagtatcatacttttgatggaaccacaagaagtcttatct	563
Db	1374	RRR	13151
Oy	564	ttaggtgacaattaactatctcttgggaagtgggaaaaaatattacaagaacgtgg	623
Db	1314	RRR	1255
Oy	624	agtattgtacctgaaggatggtgcagtgcacaatgatcatgcactccagaacaacgc	683
Db	1254	RRR	1195
Oy	684	taaaaatgctccaagaatlaactgaagagagatccatacataaagaactgtlactoca	743
Db	1194	RRR	1135
Oy	744	gggaagaatatgaccttattgtccaagtlacttgtcataltggaagaagaacttcg	803
Db	1134	RRR	10757
Oy	804	tgcgaatttcaagtgctgctgtgtctgtcgtatctcttcaatttcagctg	856
Db	1074	RRRRRRRRATCGCAAGCTCCCTGACCTGGACGCAGCAGCTCGGAATTAAATTCTG	1022

Search completed: June 28, 2002, 02:12:39
Job time: 2654 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 01:27:20 ; Search time 1914.35 Seconds
(without alignments)
10498.063 Million cell updates/sec

Title: US-09-657-631-8
Perfect score: 1489
Sequence: 1 aagtcctctctctctctgtag.....cgatctctctctctgctt 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	502.8	33.8	745	10	BG644489
2	489.8	32.9	760	10	BG584413
3	474.8	31.9	652	10	B1270380
4	465	31.2	658	10	B1263104
5	452.8	30.4	625	10	B1272922
6	451.6	30.3	651	10	B1267324
7	446.8	30.0	650	10	BG457694
8	446.2	30.0	659	10	BF639165
9	434.8	29.2	616	10	B1272802
10	433.4	29.1	664	10	B1265005
11	423	28.4	454	9	AU089497
12	419.4	28.2	659	9	BG457472
13	418	28.1	418	9	AV418181
14	413.8	27.8	663	10	B1267321
15	413.4	27.8	680	10	BF641307
16	412.2	27.7	648	10	B1265356
17	410.8	27.6	583	10	BE322878

18	410.6	27.6	660	10	BE325702
19	410.2	27.5	642	10	BG449843
20	409.6	27.5	656	10	B1265733
21	405	27.2	666	10	BG455221
22	403.6	27.1	576	10	B1272477
23	401	26.9	659	10	BF641795
24	400.2	26.9	665	10	B1266953
25	397.8	26.7	651	10	BG449019
26	397.6	26.7	556	10	BF003844
27	397	26.7	633	10	B1265090
28	396.6	26.6	603	9	BE059022
29	395.6	26.6	658	10	BF641625
30	393.8	26.4	685	10	B1271332
31	391	26.3	619	10	B1266386
32	390.6	26.3	603	10	B1266110
33	389.4	26.2	631	10	B1265759
34	386.2	25.9	655	10	B1263495
35	378.8	25.4	597	10	BG449660
36	378.6	25.4	622	10	BF639043
37	376.8	25.3	528	10	BF068331
38	375.2	25.2	648	10	B1266743
39	374.8	25.2	611	10	B1267700
40	374.2	25.1	657	9	AW584590
41	374	25.1	620	10	B1267682
42	373.6	25.1	604	9	AJ388942
43	373	25.1	624	10	BF642432
44	368.4	24.7	586	10	B1265201
45	368	24.7	643	10	BE322347

ALIGNMENTS

RESULT 1
LOCUS BG644489
DEFINITION EST506108 KV3 Medicago truncatula cDNA clone PKV3-37M24 5' end,
ACCESSION BG644489
VERSION BG644489.1 GI:13779601
KEYWORDS EST.
SOURCE Medicago.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
Medicago.
REFERENCE 1 (bases 1 to 745)
Vandenbosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D., Bowman
C.L., Craven,M.B., Cho,J. and Fraser,C.M.
ESTs from roots of Medicago truncatula 72 h after Rhizobium
inoculation, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Vandenbosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: kate@mail.bio.tamu.edu
M3928966 r1GR sequence name: MTEBP84TR More information is
available at: www.medicago.org
Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
Location/Qualifiers
1..745
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="PKV3-37M24"
/clone_11b="KV3"
/tissue_type="Seedling roots"
/dev_stage="3 days post-inoculation with Sinorhizobium
meliloti"

/lab_host="E. coli strain XL0R"
 /note="Vector: Bluescript SK-; Site.1: EcoRI; Site.2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0R cells."

BASE COUNT 239 a 132 c 157 g 217 t
 ORIGIN

Query Match 33.8%; Score 502.8; DB 10; Length 745;
 Best Local Similarity 81.8%; Pred. No. 4.1e-120;
 Matches 605; Conservative 0; Mismatches 132; Indels 3; Gaps 2;

OY 85 atgcctcctatctctctcccaatctcggaaacacattccatgaatcgttaagata 144
 DB 1 ATGCTTCATCTTCTTACTCC-ATATCTGGAAACAACATATTACTCA-CTAACGCTA 57
 OY 145 ttactcccaaaaatcaggaaccagttacatcatcgttatatttgatgctgttagc 204
 DB 58 AGATTTCGCCAAACAGAAACCAATTCCTTATGCTGTGTGTTGATGCTGTAGC 117
 OY 205 actggaagcagagtcacatgctcacaatttgatcagaacttagatctcctccgttga 264
 DB 118 ACTGGAAGCCGTCACATGTTTACCATTTTGATCAAACTTAATCTTCTCATGTTGCT 177
 OY 265 aagcaacttgatttatgatcgttaaacccggtttgattatatacgtctgaatcct 324
 DB 178 AAAATATTTGGTTTAAATTAAGAACACAGCCCGGTTGATGATACGCGATTAATCCA 237
 OY 325 gaagaagctgagagatctcctccactcttaaaagaagaagaatgtgtccgtgtg 384
 DB 238 GAATAAGCTGAAATCTTTATTCACATTTTAGACAGACAGAAAGTAGTCCGAG 297
 OY 385 agccagcaaccacaacacccgttaagcttgaggcaactgcaggtttaagcgtttgag 444
 DB 298 GATAGCGCTCCAGACACCCATTGACTTGGGCAACAGCAGCGTTTAAAGCTTTGMA 357
 OY 445 ggaagatcgtcgttaaaatatttgcaagcggtcagagatagtcagcaacaagaagtc 504
 DB 358 GGGATGCTCTGAAAATAATCTCAATCGGTAAAGGAGTTTGTTCAGCAATAGAAATGCC 417
 OY 505 cttaatttcaatcagatgcagtactatctcttgatggaaccagaaggtcttactct 564
 DB 418 TTCAATGTCACCTGATGACGTTTCTATTAATGATGGAACCAAGAAAGTTTATCTC 477
 OY 565 tgggtgacaataactatctctcttgagggaagttggaaaagaattacaagaagtcgga 624
 DB 478 TGGGTGACAGTTAACTATGCAATTTGGGAAATTTGGAAAAAATTCACAAAAACAGTGGGA 537
 OY 625 gtatgatcaggaagtggtgctgcaaatgacatatgagtcctaaggaacacagct 684
 DB 538 GTATATGATCTTGGAGGTGATCAGTTCAATGTCATATGCGATGAGTCAAAATATACGCT 597
 OY 685 aaaaatgctccaaaagtaactcgaagagagatcatatataaagaagctgtactcag 744
 DB 598 AAAATGCTCCAAAAGTTGCTGATGAGAGATCATAATTAAGAACTGTACTCAG 657
 OY 745 ggaagaataatgaccttattgacagatgaactgctgctatggaagaagaacattcgt 804
 DB 658 GGAATAAATAATGATCTCTATGTCATATGTTACTTACACTTGTGTAAGAGCATCTCGA 717
 OY 805 gcagagatttcagatgcgc 824
 DB 718 GCAGAGATTTGAAGGTAC 737

RESULT 2
 BG584413 760 bp mRNA linear EST 11-APR-2001
 LOCUS BG584413
 DEFINITION EST146173 MHAM Medicago truncatula/Glommus versiforme mixed EST

ACCESSION library cDNA clone pMHAM-15D7 5' end, mRNA sequence.
 VERSION BG584413
 KEYWORDS GI:13599477
 SOURCE EST.
 ORGANISM Medicago truncatula/Glommus versiforme mixed EST library.
 REFERENCE Medicago truncatula/Glommus versiforme mixed EST library.
 AUTHORS Eukaryota: mixed EST libraries.
 1 (bases 1 to 760)
 HARRISON, M.J., LIU, J., TOWN, C.D., VAN AKEN, S., UTTERBACK, T., CHO, J. and FRASER, C.M.
 ESTs from roots of Medicago truncatula after colonization with Glommus versiforme, 2001
 JOURNAL Unpublished (2001)
 COMMENT Contact: Harrison M.J.
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 Noble EST name: N37983e TIGR sequence name: MTDB016TK More information is available at: <http://www.medicago.org>
 Seq primer: Skm0d (CTA gaa cra gty gat cc).

FEATURES

source

1..760
 /organism="Medicago truncatula/Glommus versiforme mixed EST library"
 /cultivar="Medicago truncatula genotype Al1"
 /db_xref="taxon:119092"
 /clone_lib="pMHAM-15D7"
 /clone_lib="MHAM"
 /tissue_type="roots colonized with Glommus versiforme"
 /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glommus versiforme. The library was made from a mixture of RNA from each of these stages."
 /lab_host="E. coli strain XL0R"
 /note="Vector: Bluescript SK-; Site.1: EcoRI; Site.2: XhoI; cDNA was prepared from polyA+ enriched RNA from roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glommus versiforme. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0R cells."

BASE COUNT 241 a 143 c 153 g 223 t
 ORIGIN

Query Match 32.9%; Score 489.8; DB 10; Length 760;
 Best Local Similarity 80.6%; Pred. No. 1e-116;
 Matches 599; Conservative 0; Mismatches 137; Indels 7; Gaps 2;

OY 67 accttggtgtaagttaagtcgtctatctctccccaatctcgcgaagaacatt 126
 DB 24 ACTGTTCTACTCTTTAAATGCTCGCAATCACTTCACAAATTAAGGAAACACCTA 83
 OY 127 ctcatgagctgaagatattactcccaaaaatcaggaaccagttacatcagcgtt 186
 DB 84 CTCACTAATCGAAGATTTT-----CCAAAACAGAAACAAATTTCTCTTAAGCTGTC 137
 OY 187 atattgatcgtgtaagcactggaagcagatgctcatcatttgaactcagaacta 246
 DB 138 GTGTTTATGCTGTGACACGTGAGCGATTCATCATGTTTACCAATTTGATGACAACTTA 197
 OY 247 gatcctctcccggttgaagaacacttgatttatgattcoggttaaacccggttga 306
 DB 198 GATCTTCTCATATATGGCAAGAGTTCAGTTTATTAAGATTAACACCTGGTTTAGT 257
 OY 307 tcatagcgtcctaaccctgaagaagctgcagaaatctcgtatccactctcaagaagca 366
 DB 258 TCATATGCAAAATGATCCGGAACAGAGCTGCAAAATCTTTGATTCATCCATTACAAACAGCA 317

OY	367	gaaaatgtagtctccgctgtagccaggaacccaacacccgyltaagcttgggcaactgca	426
Db	318	gaaattgtagtccctattgattctacatcaacaacacacgattgaccttgggcaacccca	377
OY	427	ggttaagagcttctgaggggaatgctgctgaaatatatgcaagcgcgtcagggatag	486
Db	378	ggtttaagsgcttttgaattggggatgctctctgaaagatattttgcgaagcggttaaggatattg	437
OY	487	ctcagcaacagaagtgcaccttaatgcttaacatcagatgcagtatctatcttgaatgaacc	546
Db	438	ttcaccaattagaaagatccttcacattgacatttcaattgacatttgaattgattgattgaacc	497
OY	547	caagaagcttctatctcttggctgacaaatlaactcctcttggggaagtctgggaaaga	606
Db	498	caagaagcttcttattctctgctggtagacagttaacttgcatttggggaaattttagggaaaaag	557
OY	607	tttacaagaacagtgaggagtagttgatctagga-ggtggctcagtgcaaatgacatagtc	665
Db	558	tacacaaaaaagctggagatgattgatgttggagggtagatgacattgacatttgaattggcattatgac	617
OY	666	agtcctcaaggagacacagatlaaataatgctccaaagatlaacttaagagaagagatccataat	725
Db	618	agtcgcaaaagaaaacacgctaaaaatgcttccaaaaattgctgattggaagttgattcccttacct	677
OY	726	aaagaagcttctactccagggaaaagaataatgaccttattgctcagaatctacttgcgcta	785
Db	678	taagaagcttctgactcaaaagggaacaaaccatatgattcttcatgtttcacagttacttacctt	737
OY	786	tggaagagaagcattctgctgcag	808
Db	738	tggtagagaagcattctgcagctg	760

FEATURES	SOURCE	JOURNAL	COMMENT	REFERENCE	AUTHORS	TITLE	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
RESULT 3							BI270380						
							652 bp	linear	EST	18-JUL-2001			
							NEO10D1FL1094	Developing flower	Medicago truncatula	CDNA clone			
							NEO10D1FL5	, mRNA sequence.					
							BI270380						
							BI270380.1	GI:14877916					
							EST.						
							barrel medic.						
							Medicago truncatula						
							Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
							Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;						
							Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;						
							Medicago.						
							1 (bases 1 to 652)						
							Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,						
							Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.						
							Expressed Sequence Tags from the Samuel Roberts Noble Foundation						
							Medicago truncatula flower library						
							Unpublished (2001)						
							Contact: May GD						
							Plant Biology Division						
							The Samuel Roberts Noble Foundation						
							2510 Sam Noble Parkway, Ardmore, OK 73402, USA						
							Tel: 580 221 7391						
							Fax: 580 221 7380						
							Email: gdmay@noble.org						
							Insert Length: 652						
							Std Error: 0.00						
							Plate: 010						
							Row: D						
							Column: 11						
							Seq primer: TCACACAGGAACGATGATAC.						
							Location/Qualifiers						
							1..652						

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF010D11F"
/clone_1lb="Developing flower"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."

```

/note="Vector: Lambda Zap; cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

Query Match	31.9%	Score 474.8	DB 10	Length 652
Best Local Similarity	82.9%	Pred. No. 7.9e-113		
Matches 539	Conservative 0	Mismatches 111	Indels 0	Gaps 0
Oy 436	cttttgagggagaaatgcgcgtcgaataatattgcaacgcgtcagggatatgtcagcaac	495		
Db 1	CTTTTGAATGGGGAGTCTTGTAATAAATTAAGGATTTGTTACGCAAT	60		
Oy 496	agaagtcgcccttaattgttcaatcagaatgagatattctttgatgaaacccaagaagt	555		
Db 61	AGAAATACCTTCAATGTTCAACCTGATGCAGTTCTATTATGATGGAACCCAGAAAGT	120		
Oy 556	tctatcttttgggtgagacaataatactatcttcttggggaagtgggaaagaatttacaag	615		
Db 121	TCTTTTCTCTGGGTACAGTTAACTATCATTTGGGGAAATTTGGAAAAAATTTCAAAAA	180		
Oy 616	acagtgaggagtaattgtacataaggaagtgggtcagtgcaatgacataatgcagtcacaag	675		
Db 181	ACAGTGGGAGTAATGATCTTGAGAGTGGATTCAGTTCAAATGGCATATGCAAGTGTCAAG	240		
Oy 676	aacacagctaaanaatgctccaaaagcttgaaggagagatccaataaagaagctt	735		
Db 241	TATACAGCTTAAAAATGCTCCAAAAAGTTGCTGATGAGAGAATCCATTAATTAAAGAACTT	300		
Oy 736	gtaccctcagggagaanaatatgaccttatgttccacagttcttgcgtatagaagaagaa	795		
Db 301	GTATCAAAAGGGAAAAAATATGATCTCTATGTCTTAATCTTACTTACCTTTTGTAAGAA	360		
Oy 796	gcaattcgtgcagagatttcaaggctgcgtggtgtctgtcaatcccttgcatattagct	855		
Db 361	GCAATCTCAGACGAGAGATTGTAAGSTCAACATAATTTCTCCAAACCTTTGCATTTTACT	420		
Oy 856	ggccttgaatgggacatacatatctcggagcagaagataaagcttcggccccagcttca	915		
Db 421	GGATTTGATGGGACATACACATATGCTGAGAGAAATTAAGGCCAATGCCCTGCTTCT	480		
Oy 916	ggatcttaacttgaatcaatgcagaaagatagctcttaaggctcttaaaagttaatgacct	975		
Db 481	GGAGGCAATTTTAAAAATTTGAAAAAAGATAGTCTGTGGGCTCTTAAATTAATTAATCA	540		
Oy 976	tgctccatcagaatgtgaccttgggtgggatalgaaatggttgagagtggaagtgtgcaa	1035		
Db 541	TGTCCCTATTCANAATATGCACTTTTGTGGAAATTTGGATGGTGGAGGANGAATGTGACAN	600		
Oy 1036	aaaaactcttcccttaactcatcttccatcaactcctcttgaagaatgtgtg	1085		
Db 601	AGAAATACCTTTTGGCTTCACTTCTTTTACCTANTGAAGATATTGG	650		
RESULT 4				
BI263104	658 bp	mrna	linear	EST 18-JUL-2001
LOCUS	NP038611P1	NP038611P1.095	Phosphate starved	leaf
DEFINITION	clone NP038611P1	5', mRNA sequence.	Medicago	truncatula
ACCESSION	BI263104			
VERSION	BI263104.1	GI:14864204		
KEYWORDS	EST.			
SOURCE	barrelmedic.			

ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
1 (bases 1 to 658)
Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula phosphate-starved leaf library
Unpublished (2000)
JOURNAL COMMENT
Contact: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7325
Fax: 580 221 7380
Email: mjharrison@noble.org
Insert Length: 658 Std Error: 0.00
Plate: 038 Row: F Column: 11
Seq primer: TCACACGAGAAACAGCTATGAC.
Location/Qualifiers
1..658
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone_id="NF038F11PL"
/clone_lib="Phosphate starved leaf"
/tissue_type="leaf"
/dev_stage="Trifoliolate"
/note="Vector: Lambda Zap: At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20um potassium phosphate. RNA was prepared from above ground tissues."

BASE COUNT 214 a 118 c 150 g 174 t 2 others
ORIGIN

Query Match 31.2%; Score 465; DB 10; Length 658;
Best Local Similarity 83.1%; Pred. No. 2.8e-110;
Matches 528; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 295 ccgggttggtatcatcgcctgcctcaatccggaagctgcgaatctctgattccactt 354
|||||
DB 16 CCCGTTTGAGTGATCGCGGATTAATCCAGAAACATCTTTGATCCACTT 75
QY 355 ctaaaagaacagaanaatggtctcgtgagccagcaaccacacacccgttaagctt 414
|||||
DB 76 TTAGAGCAGCAGAAATGTAAGTCTCGAGATCAGCGCTCCAAAGACACCATTAAGACTT 135
QY 415 ggggcaactcaggttaaggtcttgagggagatgctgctgaaatatatgcaagcg 474
|||||
DB 136 GCGGCAACACAGGTTTAAAGCTTTGAATGGGATCTCTGAAAATAATACATCG 195
QY 475 gtaaggatagtcacagcaagaagtccttaattgctcaalcagatgcagatctatt 534
|||||
DB 196 GTAAAGGATTTGTTACCAATAGAAAGTCACTTCACTTCAACCTGATGCAAGTTTCTATT 255
QY 535 ctgatgaaacccaagaagttcttattcttgggtgagcaattactctctctggggag 594
|||||
DB 256 ATGATGAGAACCAAGAGTTCTTAATCTGGGTGACATTAATGATGATGGGAAA 315
QY 595 ttgggaaaagattccaagaagctgggagtagttgattcagaggtgggtcagtgcaa 654
|||||
DB 316 TTGGGAAAAAATTTCACAAAACAGTGGAGTATGATCTTGAGAGTGCATGATTCAA 375
QY 655 atgaacatagcagttcacaagaacacagctcaaaaatgctccaagaagctgaagag 714
|||||
DB 376 ATGGCATATTCAGTCAAGTATACAGTAAATAATGCTCCAAAAGTTGCTGATGAGAA 435
QY 715 gatccatacagaagagctgctgacccagggaagaataatgactcttattgctcagct 774
|||||

DB 436 GATCCATACATTAGAAAGCTTGACTCAAGGAAAAAATATGATCTATGTTATAGT 495
QY 775 tacttgctatggaagaagacttcctgcgaagatttcaaggtcgcgtggtctt 834
|||||
DB 496 TACTTACACTTTGTTAGAGAGCAATCTCGACCAAGATTGGAAGGTCCACACATTAATCT 555
QY 835 gctatccctgatttagcgtgcttgatgagcatatattccggagcagat 894
|||||
DB 556 CCAACCTTGATTTTATGCTGGATTTGATGGACATACATATGCTGGANNAATTTT 615
QY 895 aagctcggcccaagcttaagatctaactgaa 929
|||||
DB 616 AAGGCAATGCCCTGCTTGGAGCAATTTTAA 650

RESULT 5
BI272922
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 625)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 625 Std Error: 0.00
Plate: 091 Row: F Column: 05
Seq primer: TCACACGAGAAACAGCTATGAC.
Location/Qualifiers
1..625
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone_id="NF091F05FL"
/clone_lib="Developing flower"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."
/note="Vector: Lambda Zap: cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the uni-zap XR vector (Stratagene) and packaged using the GigaPack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant uni-zap XR vector using EXASist helper phage and the E. coli strain X1-blue MRP' (Stratagene). Excised plasmids were plated using SOUR cells."

BASE COUNT 205 a 110 c 140 g 169 t 1 others
ORIGIN

Query Match 30.4%; Score 452.8; DB 10; Length 625;
Best Local Similarity 82.7%; Pred. No. 4.2e-107;
Matches 517; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

OY		312	cgtcgcatactccatggaagaagctgcagaaactctcttgattccacttcacaaaaagaagcaaaa	371
Dd		1	CGCGATATAATCCAGAGAAGCTGCAAAATCTTTGATTCCACTTTTAGACAAAGCAAGAAAG	60
OY		372	tgtggttctcgttgaaccgaaccaaacaccccgtlaagtctggggaactgcagttt	431
Dd		61	TGTAGTTCCTGAAGATCAGCGCTCCCAAGACACCCTATGACTTGGGGCCAACGAGGTTTT	120
OY		432	aagccttttggagggaatgtctgtcgtaaatatatattgcaagcgcacaggatatagtcag	491
Dd		121	AAGCTTTTGAATGGGGATGCTTCTGAAAAAATACGTGCAATCGTAAGGGATTTTGTAG	180
OY		492	caacaagaagtcgccctaattgtccaacagatgcagatcatctatctctgtatgtaaaccaaga	551
Dd		181	CAATAGAGATACCTTCATATGTTCAACCTGATGCAATTCTATTTATGATGGAACCCAAGN	240
OY		552	aggtcttatctcttgggtgtagcaaatbaactatctctctgggaagltgggaaaaagaattac	611
Dd		241	AGGTTCTTATCTCTGGGGTGCAGATTAACTATGACATTTGGGAAATTGGGAAAAAATTCAC	300
OY		612	aaagaacgtgggaagtgaactgatctcagagggtgggtcgaatgcgaataatgcagcttc	671
Dd		301	AAAAACAATGGAGATTAATGATCTTGGAGGTGCATCATTTCAAATGGCATGACAGTGTG	360
OY		672	aaggaaccagcctaanaatgtctccaaaagtaacctgaagagagagatgccataaagaa	731
Dd		361	AAAGATATACACTTAAAAATGCTCCAAAAGTTGCTGATGAGAGATGCATCATTAAGAA	420
OY		732	gcttgtlactccaaggaaaataatgatcctttaattgtcacagtlacttgcgtlagaag	791
Dd		421	GCTGTACTCAAGGGAAAAAATATGATCTCTATGTTATGTTACTTACACTTTGGTAG	480
OY		792	agaagcatttcgttcagaagatttcaaagtcgcgtgtgtctgtctgaatcccttgattt	851
Dd		481	AGAAGCATCTGACAGACAGATTTTGAAGGTACACACATATTTCTCCCAACCTTGCATTTT	540
OY		852	agctggtcttgatcgaggcataacatactccggaagcagagataaagtcctgcggcccaag	911
Dd		541	AACGTGATTGATGGACATACACATTAATGCTGGAGAAAAATTAAGGCCAATGCCCTGAC	600
OY		912	ttcaggaatcaactgaatcaatgc	936
Dd		601	TTCTGGAGCCAATTTTAAAAAATGC	625
RESULT		6		
B1267324				
LOCUS		B1267324	651 bp	mRNA linear EST 18-JUL-2001
DEFINITION		NF105B1ZIN1P101 Insect herbivory	Medicago truncatula cDNA clone	
ACCESSION		NF105B1ZIN 5,	mRNA sequence.	
VERSION		B1267324		
KEYWORDS		B1267324.1 GI:14872249		
SOURCE		EST.		
ORGANISM		barrel medic.		
		Medicago truncatula		
		Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;		
		Rosidae; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;		
		Medicago.		
REFERENCE		1 (bases 1 to 651)		
AUTHORS		Korth,K., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Florese,		
TITLE		H.R., Imman,J.T., Weller,J.W. and May,G.D.		
JOURNAL		Expressed Sequence Tags from the Samuel Roberts Noble Foundation		
COMMENT		Medicago truncatula Insect herbivory library		
		Contact: Korth K		
		Dept. of Plant Pathology		
		University of Arkansas		
		217 Plant Science Building, Fayetteville, AR 72701, USA		
		Tel: 501 575 5191		
		Fax: 501 575 7601		
		Email: khorth@comp.uark.edu		

```
Insert Length: 651      Std Error: 0.000
Plate: 105 row: B      column: 12
Seq primer: TCACACAGGAACAGCTATGAC.
```

FEATURES	Location/Qualifiers
source	1. .651

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF105B12IN"
/clone_lib="insect herbivory"
/tissue_type="local and systemic leaves"

```

```

/ode_stage="mature"
/node="Vector: Lambda zap: Library was produced from fully
expanded M. truncatula leaves of plants fed upon by
Spodoptera exigua (beet armyworm) for 24 hours. Systemic
(undamaged leaves from injured plants) and wounded leaves
were harvested and pooled "
114.0 142.0
208

```

BASE COUNT	200	a	114	c	142	g	107
ORIGIN							

Query Match	30.3%	Score 451.6	DB 10	length 651
Best Local Similarity	83.3%	Pred. No. 8.7e-107		
Matches 525	Conservative	0	Mismatches 104	Indels 1
				Gaps 1

[illegible]

22 AATAACAAGAACCAATTCCCTTATGCTGTGCTGTTCATGCTGCTAGCACCCTGGAACC 81
Db
315 gaattccattatctacaattttgaatcagaacttagatctccttcccatgaaaaagaactta 274
Ov

Db 82 GTGTCACATGTTTACCATTTTGATCAAACTTAATCTTCTTCATGTGTGTAACAGATGTC 141

275 agtttatgattcgtttaaccggtttgagttcatacgtcgtctaatacctgaagaagctg 334

Dd 142 AGTTTATATAAGACACGCGGGTTTGATGCATACGCGGATATCCAGAAGACTG 201

Ov 335 CAGAACTCATTATTCGCACTCTGTAAAAAAGAAAAGAAAAATATTAATTCCTCTTGAAGCCAGCAAC 394

202 CAAATCTTGATTCACCTTTTAGACGACGAAAGCTAGTTCCTGAGGATCAGCGCT 261

Q7 395 ccacacacaccgttaagcttggggcaactgcaggttcaagccttggagggggaatgctg 454

Db 262 CCAAGACACCCATTAGACTTGGGGCAACAGAGCTTTAAAGCITTTTGAATGGGGATGCTT 321

Db

322 CTGAAAAAATACCTGCATCGGTAGGCGATTGTTCACGCAATAGAGTACCCTCAATGTTCC 381

515 aatcagatgcagltatctatctcttgatcgaaacccaagaaggtcttatactttggatgacaa 574

Db 382 AACCTGATGCAGTTCTATTATTGATGGAACCCAGAAGTCTTATCTCTGGGTGACAG 441

Db

	5/5	ttcaatcatctcttgggaaagtcttggaataaagaattacaagacagtcgtggaatgatc	b34
OY			
	442	TTTAACTATCATTTGGGGAATTGGGCAAAAAAATTCACAAAACAGTGCGAGTAATGCATC	501

635 tagagagtcggtcagtcgaatatgcaagtcctcaaggaacacagctaaaatgttc 694

Db 502 TTGGAGGTGATCAGTTCAATGCGCATATGCGAGTGTCAAGGTATACAGCTAAATAATGCTC 561

695 caaaagtaactgaagagagataccctacatataagaagcttgaacctccgggaagaanaat 754
 0y |||||
 Db 562 CAAAGTCTGATGAGAGAGATCCCTTACATTAAAGAGCTTGACTCAAGGGAAAAAAT 621
 |||||

QY 755 atgaccttatgttcacagtacttcgct 784 .

Db 622 ATGATCTCTATGTCATAGTACTTACCT 651

RESULT 7
BG457694

LOCUS	650 bp	mRNA	linear	EST 19-MAR-2001
EG457664				
MF106EL1PL1F1085	Phosphate starved	leaf	Medicago truncatula	CDNA

ACCESSION	clone NF106E11PL 5', mRNA sequence.					
VERSION	BG457694					
KEYWORDS	BG457694.1 GI:13381019					
SOURCE	EST.					
ORGANISM	barrel medic. Medicago truncatula Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.					
REFERENCE	1 (bases 1 to 650)					
AUTHORS	Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores, H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.					
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula phosphate-starved leaf library					
JOURNAL	Unpublished (2000)					
COMMENT	Contract: Harrison MJ Plant Biology Division The Samuel Roberts Noble Foundation 2510 Sam Noble Parkway, Ardmore, OK 73402, USA Tel: 580 221 7325 Fax: 580 221 7380 Email: mjharrison@noble.org Insert Length: 650 Std Error: 0.00 Plate: 106 Row: E Column: 11 Seq primer: TCACAGCAAAACACTATGAC. Location/Qualifiers 1..650 /organism="Medicago truncatula" /db_xref="taxon:3880" /clone="NF106E11PL" /clone_lib="Phosphate starved leaf" /tissue_type="leaf" /dev_stage="trifoliolate" /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoagland's solution containing only 20um potassium phosphate. RNA was prepared from above ground tissues."					
BASE COUNT	211 a 110 c 153 g 176 t					
ORIGIN						
Query Match	30.0%; Score 446.8; DB 10; Length 650;					
Best Local Similarity	81.5%; Pred. NO. 1.6e-105;					
Matches 517; Conservative	0; Mismatches 117; Indels 0; Gaps 0;					
OY	265 aacgaacttgatttaagatcggtaaaccggttgagtgcatacgctcaatccct 324					
DB						
	11 AACGATGTGGATTTTAATAAAGAACAACACCCTTGTAATGCATACCGCATTAATCCA 70					
OY	325 gaagaagctgcagaatctctgatccaccttcataaagaacagaaaatlygttcctgtg 384					
DB						
	71 AAGGAAGCAGCAGAGTCTCTGATTCACATTTTAAGCAAGCAAGAAAGTGTCCTGTT 130					
OY	385 agccaagcaaccaacaacaccggttaagcttggggcaactgaggtttaaggcttttggag 444					
DB						
	131 AATCGCAACCCCAAACAACCCGTTAAAGCTTGGGCCAACAGCCGCTTTAAAGGCTTTTGAT 190					
OY	445 gggaagctcgtctgaataatatatgtgcaaagcggtcaggagatatgtccagcaacaagatgcc 504					
DB						
	191 GGAAATTTCTCCGAATGATATTGGAAGCGGTGTCGAGTTTGCTCAAAAAGAGAACGACG 250					
OY	505 cttaagtccaatcaagatgcagatctcatctatcttgatgaagcccaagaaggtcttatct 564					
DB						
	251 TTTAACGTACATATAGATGCAAGTAGAGAAATTAATGATGTCACAAAGAGGTTCTTATTTA 310					
OY	565 ttgggtgaacattaaactatctcttgggggaagttgggaaaaaatltaacaaagacagtygga 624					
DB						
	311 TGGGTGCAATTAATCATATGTTTTGGGAAACTTGGGAAAGATTTTTCAGAGACAGTGGCA 370					
OY	625 gttagtgatctaggaggttgggtcagatgcaaatgacatgacatgcagttctaaggaacacagct 684					

Db	Accession	Length	Score	E-value	Organism	Source	Comment
Db	371	GTACCTGATCTTGGAGGGGATCAGTTCAATGGTATATGACGTCTCAGAGAAACA	430				
Qy	685	aaatctctcaaaagtaactgaagagagagatccatacctaagaagctgttaccgcg	744				
Db	431	AAAAAGCTCCACAAAGTACTCTCAGGGAGAGGATCCATACATTAATAAGATTGTACTTAA	490				
Qy	745	ggaagaataatgaaccttaatgtcttcacagttactgcgcgtatggaagaagaagcttcgt	804				
Db	491	GGAAGAAATATTAACCTCTATGTTACACAGTATTTATTCGGGTTTGCAAGAGCATCTCGT	550				
Qy	805	gcagagatttcaagatcgcgtgtgtgttcgcctaactccttgatcttaactgtcttgcg	864				
Db	551	GCTGAATTTTGAAGTACTACTAATGATTTCTCCTAATCTGTGCACTTTAGCTGGCTATCAT	610				
Qy	865	ggggcatatcatatcttcggagagagatlaag	898				
Db	611	GGCACAATATCATATTTCAGAGAGAGATTAAG	644				
RESULT	8						
LOCUS	BF639165	659 bp	mRNA	linear	EST 19-DEC-2000		
DEFINITION	BF639165	BF639165	BF639165	BF639165	BF639165	BF639165	BF639165
ACCESSION	BF639165	BF639165	BF639165	BF639165	BF639165	BF639165	BF639165
VERSION	BF639165.1	GI:11903323					
KEYWORDS	EST.						
SOURCE	barrel medic.						
ORGANISM	Medicago truncatula						
REFERENCE	1 (bases 1 to 659)						
AUTHORS	Lin,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores						
TITLE	H.R., Imman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.						
JOURNAL	Expressed Sequence Tags from the Samuel Roberts Noble Foundation						
COMMENT	Medicago truncatula phosphate-starved leaf library						
FEATURES	Unpublished (2000)						
source	Contract: Harrison MJ						
	Plant Biology Division						
	The Samuel Roberts Noble Foundation						
	2510 Sam Noble Parkway, Ardmore, OK 73402, USA						
	Tel: 580 221 7325						
	Fax: 580 221 7380						
	Email: mjharrison@noble.org						
	Insert Length: 659						
	Std Error: 0.00						
	Plate: 094						
	row: A						
	column: 12						
	Seq primer: TCACACAGAAACAGCATATGAC.						
	Location/Qualifiers						
	1..659						
	/organism="Medicago truncatula"						
	/db_xref="taxon:3880"						
	/clone="NF094A12PL"						
	/clone_lib="phosphate starved leaf"						
	/tissue_type="leaf"						
	/dev_stage="trifoliolate"						
	/note="Vector: Lambda Zap; At the trifoliolate stage, M.						
	truncatula plants were transplanted to phosphate-free sand						
	and grown for a further 30 days. During this 30 day						
	period, the plants were fertilized twice weekly with 1/2						
	Hoaglands solution containing only 20uM potassium						
	phosphate. RNA was prepared from above ground tissues."						
BASE COUNT	208 a	114 c	151 g	180 t	6 others		
ORIGIN							
Query Match	30.0%	Score 446.2	DB 10	Length 659			
Best Local Similarity	80.1%	Pred. No. 2.2e-105					
Matches 520	Conservative 0	Mismatches 129	Indels 0	Gaps 0			
265	aacgaacttggtttatgattcgtttaaccggttgattcatatcagctgataacccct	324					

```

Db 11 AACGATGTTGATTTTATTAATAGACAACACCCGTTTGATGCAATACGAGTAATACCA 70
Qy 325 gaagaagctgcagaatctctgattccactcttaaaagaagagaatggtgtctctg 384
Db 71 AAGGACAGCAGAGCTCTGATTCACCTTTTACAGCAACGAGAAAGGTTCTCTGT 130
Qy 385 agccagcaaccacacacccggttaagcttgaggcaactgcaggtttaaggtcttgag 444
Db 131 AATTCGACCCCAAAACACCGTTAGCTTGAGGCAACGCGGTTTAAAGCCTTTGGAT 190
Qy 445 ggaagctgctggaataatataatgcaagcgtgagagatagctcagcaagaagatgcc 504
Db 191 GGAATTTCTCCGAAATGATATTTGGAAGCGGTGCGAGTTTGCTCAAAAAGAGACAGC 250
Qy 505 cttaattcaatcagatgcagtactctatctctgtagaagcaagaggttctatctt 564
Db 251 TTTACGTAACATCAGATGAGTAGAGAAATTAATGATGATACACAAAGGTTCTTATTTA 310
Qy 565 tgggtgcaattactatctctcttgaggaaqtlgggaaagaatttacaagaagacagtgga 624
Db 311 TGGGTGCAATTAACCTATGTTTGGAAACTTGGAAAGATTTTTCAGACAGAGTGCA 370
Qy 625 gtaqgtgactgaagggtgggtcagtgcaaatgacataatgcagctcagaaagacagct 684
Db 371 GTAGCTGATCTTGGAGGGGATCAGTTCAATGATATATGAGCTCAGAGAACAGCA 430
Qy 685 aaaaatgctccaaaagactctgaaagagagatccatacataaagaagctgtgactcag 744
Db 431 AAAAACTCTCCACAAAGTACCTCAGGAGAGATGTCATTAATAAAAGATTGTACTTAA 490
Qy 745 ggaagaataatgaccttaattgttcacagttacttgcgtacatgaaagaagacattcgt 804
Db 491 GGAAGAATATTTACCTCTATGTTTCACAGTTATTTGGCGTTTGGCAANAAAGATCTCGT 550
Qy 805 gcagagatttcaaggtcgtggtgtctgtaacatcccttgacatttagctggtctgag 864
Db 551 GCTAAATTTTGAAGAGTACATTAATGTTCTCTCAATCCTTGCAATTTTANCTGCTATCAT 610
Qy 865 ggagcatatacatatcttcgagagcagatataagttctcgccagctt 913
Db 611 GGCACTATATNCATATTCAGAGAGAGATNTNAGGCAATTTCTCTCTGCTT 659

```

```

RESULT 9
LOCUS BI272902 616 bp mRNA linear EST 18-JUL-2001
DEFINITION NF098609FL1071 Developing flower Medicago truncatula cDNA clone
ACCESSION BI272902
VERSION BI272902.1 GI:14882630
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
Medicago.
REFERENCE 1 (bases 1 to 616)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Imman, J.T., Weller, J.W., and May, G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation.
JOURNAL Medicago truncatula flower library
COMMENT Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 616 Std Error: 0.00
Plate: 098 row: 6 column: 09

```

```

FEATURES             Seq primer: TCACACAGGAACAGCTATGAC.
source               location/Qualifiers
1..616
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF098609FL"
/issue_lib="Developing flower"
/issue_type="Developing flower"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/notes="Vector: Lambda Zap; cDNA was prepared from polyA+
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRF' (Stratagene). Excised plasmids were plated
using SOLAR cells."
BASE COUNT          194 a          96 c          137 g          187 t          2 others
ORIGIN
Query Match          29 2%; Score 434.8; DB 10; Length 616;
Best Local Similarity 81.5%; Pred. NO. 2e-102;
Matches 502; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
Qy 452 ctgctgaaaataatataatgcaagcgtgacagagatagctcgaacacagaagtgcccttaatg 511
Db 1 CTCTGTAAGAAAGATATCTCAATTCGATAGCGTATGATTTGTTGACCAACAGAAATACCTTCAATG 60
Qy 512 ttaacatcagatgcagatctatctcttgaatgaagcaagagaggttcttacttcttggtga 571
Db 61 TTCAACCTGATGCGAGTTTCTATTAATGATGAGAACCCAAAGAGGTTCTTATCTATGTTGTTGA 120
Qy 572 caatlaactatctctcttgaggaaqtlgggaaagaatttacaagaagacagtgaggatgtg 631
Db 121 CAGTTAACTATGCAATTTGGGACTTTGGGAAAAAATTCACAAAACAGTGGAGTAATGG 180
Qy 632 atcagaaggtgggtcagtgcaaatgacataatgcagctcgaagaaacacagctcaaaatg 691
Db 181 ATCTGGAGGTGATCGCTTCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 240
Qy 692 ctccaaaagtaacctgaagagagagatccatacataaagaagctgtgactcccaaggaaaga 751
Db 241 CTCCAAAAAGTTGGCTGATGAGATGATGATATTAAGAGCTTGTCTCAAGGAGAAA 300
Qy 752 aataatgaccttattgttcacagttacttgcgtacatgaaagaagacattcgtgagagaga 811
Db 301 AATATGATCTCTATGTTCAATGTTACTTACACTTGTGTAAGAGCAATCTCGGCTGAGA 360
Qy 812 ttctcaagctgcgtggtgtgtctgtaactcttgacattttagcttggtcttgatggagagat 871
Db 361 TTTTGAAGTGCACATCAATTTCTCTCAACCTTGCAATTTAGTGGATTTGAGGAGCAT 420
Qy 872 atacataatccgagaagagatataaggtctcgccccaagcttcaagatctaattgaatc 931
Db 421 ACAATATGCTGTGGAAGAAATTAAGGCCCAATGCCCTTCTGGAGCCAGTTTAAAA 480
Qy 932 aatgcagaaagatagccttcaagctttaaagtgtaatgacacgttgcctatcaagaatt 991
Db 481 AATGCAAAAGATATGTTCAATCAGGCTTTAATATGAATATTCATGCTTATCAAAAT 540
Qy 992 gcaatttggtgagataggaatggtgagagtggtgagtggtggtggtggtggtggtggtggt 1051
Db 541 GCACTTTTGGTGAATTTGGAATGCTGAGAGAAAGAAANTGAGACANMAAACTTTTCTGCTG 600
Qy 1052 ctcaatcttctatata 1067
Db 601 CTTCATTTTCTCTTTTA 616

```

OY	353	tctcaaaagaagagaanaatggttctccttgtagccagaaccaacaccggttaagc	412
Dd	297	TTTTACAGCAGACGAAAGAAGTGTTCTTCCTGTATATCGAACCCAAAACCCGTTAAGC	356
OY	413	ttagggacaactgcagglttaaggctttggagggaatgcgtctaataattgcaag	472
Dd	357	TTGGGGCAACAGCGCGGTTTTAAGCCTTTGGATGGAAATTCTCCGAATTGATTGGGAAG	416
OY	473	cgttcaggatatgtcccacaacagaagtgccttatattcatatgatgaatctta	532
Dd	417	CGGTGTCGAGTTTGCTCAAAGAGAGACAAGCTTTTAACGTACATCATGATGCAGTAGGAA	476
OY	533	tctcttagtgaagccccagaaggltcttaactttgggtgagcaactiaactctcttgagg	592
Dd	477	TATATGATGTCACACAAAGAAGGTTCTTATTTATGGGTGACAAATTACTATVGTTTGGGAA	536
OY	593	agttggaaaagaatttaaacaaagacagtgtaggtatgtatctlaagagggtgtagtc	652
Dd	537	ACTGTGGGAAAAATTTTTTCAGAGACAGTGTCAGTACTCTATCTTGGAGGGGATCATGCC	596
OY	653	aatgacatatgcagctctcaaggaacacagcctaaaatgtctcaaaagctcgaagag	712
Dd	597	AATGTATATTCAGTCTCAAGAGACAAAGCAAAAAAGCTTCACACAGTACCTCAGGAG	656
OY	713	aggatcc	719
Dd	657	ANGATCC	663
RESULT	11		
AU089497			
LOCUS	AU089497	454 bp	mRNA linear EST 23-AUG-2000
DEFINITION	AU089497 Lotus japonicus flower bud cDNA Lotus japonicus cDNA clone		
ACCESSION	Ljfb-u31, mRNA sequence.		
VERSION	AU089497		
KEYWORDS	AU089497.1 GI:9279021		
SOURCE	EST.		
ORGANISM	Lotus japonicus.		
JOURNAL	Lotus japonicus		
MEDLINE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
COMMENT	Spermatophytas; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;		
	Lotus.		
REFERENCE	1 (bases 1 to 454)		
AUTHORS	Endo,M., Kokubun,T., Takahata,T., Higashitani,A., Tabata,S. and Watanabe,M.		
TITLE	Analysis of expressed sequence tags of flower buds in Lotus japonicus		
JOURNAL	DNA Res. 7, 213-216 (2000)		
MEDLINE	20363098		
COMMENT	Contact: Masao Watanabe Fac.Of Agri. Iwate University 3-18-8 ueda, Morioka, Iwate 020-8550, Japan Tel: 81-19-621-6152 Fax: 81-19-621-6177 Email: nabe@iwate-u.ac.jp This clone was obtained from a 5'end. Location/Qualifiers 1..454 /organism="Lotus japonicus" /db_xref="taxon:34305" /clone_id="Ljfb-u31" /clone_1id="Lotus japonicus flower bud cDNA" /tissue_type="mature flower bud" /lab_host="SOLR" /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2: XhoI; variety-Gifu B-129; Plants of Lotus japonicus Gifu B-129 were grown in a greenhouse condition. Sepal was removed from flower bud. The mRNA was isolated from flower bud of 0.7-1.0cm in length just before anther dehiscence."		
FEATURES	source		
BASE COUNT	130 a	101 c	89 g 133 t others
ORIGIN			

Query Match 28.4%: Score 423; DB 9; Length 454;
 Best Local Similarity 99.3%: Pred. No. 2.2e-99;
 Matches 445; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

OY 24 gtgtcattgactgaagcatgactcttctaataatgcatcactgactggttcatgtt 83
 Db 9 GTTGCAATGGAGCTAAAGCCATGACCTTCTTAATTAATCTCATGACCTTGTGTCATGTT 68
 OY 84 aatgctgtatctcttctccccaatatactcgaaacaacatctcatgaatcgttaagat 143
 Db 69 AATGCTGTATCTCTCTCCCAATATCTCGAAACAACATCTCATGAATCGTAAGAT 128
 OY 144 attaccctcccaaaatcaggaaccagttatcatatagcgtgtatatttgatgctgtag 203
 Db 129 ATTACTCCCCAAAATCAGGAACCACTTACATCATACGCTGTATATTTGATGCTGTG 188
 OY 204 cactggaagcagaagtcactgctctacaaatttgatcagaacttagatctctccgttga 263
 Db 189 CACTGGAAGCAGATCCATGCTCTCAATTTTGATCAGAACTTGATCTCCTTCCGTTGA 248
 OY 264 aaacgaacttgagtttattagattcgggttaaacccggttgagttcataagcgtatacc 323
 Db 249 AAACGAACCTTGAGTTTATGATTGGTTAAACCGGTTTGAGTTCAATGAGTCTTAATCC 308
 OY 324 tgaagaagctgcaagatctctgactcttaagaagaagcagaagaatggttctctt 383
 Db 309 TGAAGAA-CTGCAGATCTCTGATTCACCTTCTAAAGAA-CAGAAATGTGTTCTCTGT 366
 OY 384 gtagcgaacaccccaacacacccggttaagcttgaggcaactgcaagtttaagcgtttga 443
 Db 367 GAGCCANCAACCCACACACCCGTTAAGCTTGAGGCACTGAGTTTAAAGCTTTTGA 426
 OY 444 ggggaatgctgctggaataatattatgcaa 471
 Db 427 GGGGAATGCTGTGAATAATATATTGCCA 454

RESULT 12

BG457472

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

659 bp mRNA linear EST 19-MAR-2001
 NF106A03PL1F1019 phosphate starved leaf Medicago truncatula cDNA
 clone NF106A03PL 5', mRNA sequence.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 659)
 Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
 H.R., Imman,J.T., Weiler,J.W., May,G.D. and Harrison,M.J.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula phosphate-starved leaf library
 Unpublished (2000)
 Contact: Harrison MJ
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7325
 Fax: 580 221 7380
 Email: mharrison@noble.org
 Insert Length: 659 Std Error: 0.00
 Plate: 106 row: A column: 03
 Seq primer: TCACACAGGAACACGCTATGAC.
 Location/Qualifiers
 1..659

FEATURES

source

/organism="Medicago truncatula"
 /db_xref="taxon:3880"

/clone="NF106A03PL"
 /clone.lib="phosphate starved leaf"
 /tissue.type="leaf"
 /dev.stage="trifoliolate"
 /note="Vector: Lambda zap. At the trifoliolate stage, M.
 truncatula plants were transplanted to phosphate-free sand
 and grown for a further 30 days. During this 30 day
 period, the plants were fertilized twice weekly with 1/2
 Hoaglands solution containing only 20mM potassium
 phosphate. RNA was prepared from above ground tissues."
 BASE COUNT 205 a 115 c 145 g 193 t others

Query Match 28.2%: Score 419.4; DB 10; Length 659;
 Best Local Similarity 78.8%: Pred. No. 2.1e-98;
 Matches 514; Conservative 0; Mismatches 132; Indels 6; Gaps 1;

OY 557 ctatctctgggtgacaaatgaactatctcttgaggagtggaagaaagattacaaga 616
 Db 8 CTATCTCTGGGTGACAGTTAACTATGCAATGGGAAATTTGGAAATAATTCACAAAAA 67
 OY 617 caatggagtagttgactcagaagtggtgcagtgcaaatgacatatgactcagaaga 676
 Db 68 CAGTGGAGCTAATGATCTTGAGGTGATCAGTTCAATGAGCATATGCAATGCTCAAACT 127
 OY 677 aacagctaaagaatgtctcaaaagttactcgaagagagagatcatatcataaagaagcttg 736
 Db 128 ATACAGCTAAATATGCTCCAAAAGTTGCTGATGAGAGAGATCATATCAATTAAGAACTTG 187
 OY 737 tactccagggaagaatataatgacattatgctacagttactgctgctatggaagaaga 796
 Db 188 TACTCAAGGAAATAATATGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 247
 OY 797 caattgagcagagatttcaaggtcgcgtggtgctgctatccttgcattttagctg 856
 Db 248 CATCTGACACAGATTTTGAAGGTGATCAGATTAATCTCCAAACCTTCATTATTAAGCTG 307
 OY 857 gcttgaatgggcatatataatctccggagcagagataggtctcggcccaagcttag 916
 Db 308 GATTTGATGGACATACATATGCTGAGAGAAATTTAAGGCCAATGCCCCGCTTCTG 367
 OY 917 gatcctaactgaatcagcagaagaatagctcttaagcgtcttaagatgacacatt 976
 Db 368 GAGCCAAATTTTAAAAAATGCAAAAAGATAGTGTGAGGCTCTTAATTAATTAATCAAT 427
 OY 977 gtccctataagaatgacatttgggtggaatgaaatgggtggaatggaatggaatggaat 1036
 Db 428 GTCCCTATCAGAAATTCACCTTTTGGTGAATTTGAGTGTGAGGAGGAGGAGGAGGAGG 487
 OY 1037 aaatcttctctactatctcttctatctctctcctcgaagatggttgatcttgta 1096
 Db 488 GAATACTTTTGTCTGCTTATCTTTTACTTACTTAAAGATATGATGCTGTGACCC 547
 OY 1097 ataacc-----caatggcaaaatcgtcagttatggaagctgagcgaataactag 1150
 Db 548 CAATATCATTATATATGTCACACTTCGTCCTGTGATCTTGAGATGAGCAATCAAG 607
 OY 1151 ctgttaaaacaatcttgagatgcaaaatccaacaatccagatccttata 1202
 Db 608 CTGTACATTAAATCTTGANAGGCAAAATCAATATACACTTCTTCTGGA 659

RESULT 13

AV418181

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

418 bp mRNA linear EST 23-MAY-2000
 AV418181 Lotus japonicus young plants (two-week old) Lotus
 japonicus cDNA clone M1153g11_r 5', mRNA sequence.
 AV418181
 AV418181.1 GI:7747359
 EST.
 Lotus japonicus.
 Lotus japonicus

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
 Lotus.
 1 (bases 1 to 418)
 AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
 TITLE Generation of 7137 non-redundant expressed sequence tags from a
 legume, Lotus japonicus
 JOURNAL DNA Res. 7 (2), 127-130 (2000)
 MEDLINE 20277479
 COMMENT Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1. 418
 source
 /organism="Lotus japonicus"
 /db_xref="taxon:34305"
 /clone="MM153j11_r"
 /clone_lib="Lotus japonicus young plants (two-week old)"
 /dev_stage="young plants (two-week old)"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI; isolate=Mykojima Mg-20"
 BASE COUNT 121 a 61 c 99 g 117 t
 ORIGIN
 Query Match 28.1%; Score 418; DB 9; Length 418;
 Best Local Similarity 100.0%; Pred. No. 4.4e-98;
 Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 189 attgatgctgtgtaacctggaagcagatgcacatgctacatttgatcagaactaga 248
 |||||||
 Db 1 ATTGATGCTGTGTAAGCACTGGAAGACAGATGCTACATTTTGATGCAACTTGA 60
 QY 249 tctctcccgctgtaaaacgaactgattatgatcggttaaacccggttgagtc 308
 |||||||
 Db 61 TCTCTCCCGCTGTAAACGAACTTGATTGATTCGGTTAAACCCGCTTGAAGTTC 120
 QY 309 atagcgtcattcctgaagaagctgcaagatctctgattccactctctaaagaagaga 368
 |||||||
 Db 121 ATAGCTGCTATCTCTGAAGAGCTGCAGATCTCTGATTCACCTTCTAAAGAGAGA 180
 QY 369 aaatggttcctgtagcagcaaccacaacccggttaagcttggggcaactgag 428
 |||||||
 Db 181 AAATGTTGCTCTGTGAGCCAGCAACCAACACACCCGTTAAAGCTGGGCACTGAGG 240
 QY 429 tttaagccttttgaggagggaatgctgctgaaataatattgcaagcggtcaggatgct 488
 |||||||
 Db 241 TTTAAGCTTTTGGAGGGGAATGCTGTGAAAAATATTTGCAAGCGGTCAAGGATATGCT 300
 QY 489 cagcaacaagaagtgccttaagtcaatcagatgcagatctctattcttgagaacca 548
 |||||||
 Db 301 CAGCAACAAGAGTGCCTTATGTTCAATCAGATGCAATCTATTCTTGATGGAACCA 360
 QY 549 agaaagttctatcttgggtgacaatactatcttggggaagttgggaaaga 606
 |||||||
 Db 361 AGAAGTCTTATCTTTGGGTGACATTAATATCTCTGGGGAAGTTGGAAAAAGA 418
 RESULT 14
 BI267321
 LOCUS BI267321 663 bp mRNA linear EST 18-JUL-2001
 DEFINITION NF105B10INJF1089 Insect herbivory Medicago truncatula cDNA clone
 NF105B10IN 5', mRNA sequence.
 ACCESSION BI267321
 VERSION BI267321.1 GI:14872243
 KEYWORDS EST.
 SOURCE baitrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 663)
 AUTHORS Korth,K., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
 ,H.R., Inman,J.T., Weller,J.W. and May,G.D.
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula insect herbivory library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Korth K
 Dept. of Plant Pathology
 University of Arkansas
 217 Plant Science Building, Fayetteville, AR 72701, USA
 Tel: 501 575 5191
 Fax: 501 575 7601
 Email: kthorh@comp.uark.edu
 Insert length: 663 Std Error: 0.00
 Plate: 105 row: B column: 10
 Seq primer: TCACACAGAAACAGCTATGAC.
 Location/Qualifiers
 1. 663
 source
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF105B10IN"
 /clone_lib="Insect herbivory"
 /tissue_type="local and systemic leaves"
 /dev_stage="mature"
 /note="Vector: Lambda Zap. Library was produced from fully
 expanded M. truncatula leaves of plants fed upon by
 Spodoptera exigu (beet armyworm) for 24 hours. Systemic
 (undamaged leaves from injured plants) and wounded leaves
 were harvested and pooled."
 BASE COUNT 205 a 132 c 134 g 188 t 4 others
 ORIGIN
 Query Match 27.8%; Score 413.8; DB 10; Length 663;
 Best Local Similarity 78.4%; Pred. No. 6.2e-97;
 Matches 520; Conservative 0; Mismatches 136; Indels 7; Gaps 2;
 QY 34 actaaagcagactctctaatagatcctatgacaccttggctgtaagtgcgtc 93
 |||||||
 Db 7 ACACAAAGCATGAGATGTTCTTAATTAACATCATATTTACTCTTCTTGAGCCGTACA 66
 QY 94 atctctctcccaataatctcgaaacaacatctcatgaatcgttaagatattactcccc 153
 |||||||
 Db 67 ATCTCTCTCCCAATATCTTGGAAACAATATCTACCAATGCTTAAATATTTCT----- 121
 QY 154 aaaatcagaagaccagttatcatatagctggtatatttgatgctggtgacatggaagc 213
 |||||||
 Db 122 -CAAAACAAAGAACTTAACCTGTAAGCGTGTGCTTTGATGCTGTGACGACCGGTAGC 180
 QY 214 agagtcacatgctacaaatttgatcagaactgatctccctccgctgtaaaacgaact 273
 |||||||
 Db 181 CGGTTCATGATCTACCATTTTGAATGATCAAGAACTTGTGATCTTTCATGCAACGATGTT 240
 QY 274 gagtttaagtatcgtttaaaccgggttgagttcaatcagctgtaacccctgaagaact 333
 |||||||
 Db 241 GAGTTTATTAATAAGCAACACCCGTTGAGGCTACGACGATTAATCAAGGAAGCA 300
 QY 334 gcaagaatctgattccactcttaaaagaagcagaagaatggttctctgtgagccagca 393
 |||||||
 Db 301 GCAGAGTCTCTGATTCACATTTTGAAGCAAGAAAGATGCTTCTGTATATCTGCAA 360
 QY 394 cccaacaccccggttaagcttggggaactcagctgagtttaagcgtttgaggaggatgct 453
 |||||||
 Db 361 CCCAAACACCCGTTAAGCTTG66GCAACAGCCGGTTTAAAGCTTTGAATGGAATTTCT 420
 QY 454 gctgaaataatattgcaagcgtgcaaggatgctcagaacaagaagtgccttaagt 513
 |||||||
 Db 421 TCCGAATTTGATTTGGAAGCGGTGTCGAGTTTCTCAAAAAGAGAACAGCTTTAACGTA 480
 QY 514 caatcagatgcagatctattcttgatggaaccagaaggttctatctttgggtgac 573
 |||||||

[illegible]

Search completed: June 28, 2002, 02:11:19
Job time: 2639 sec

THIS PAGE BLANK (USPTO)
